

The British Journal for the Philosophy of Science

VOLUME XI

FEBRUARY, 1961

No. 44

A METHODOLOGICAL PROBLEM IN RHEOLOGY *

A. GRAHAM, G. W. SCOTT BLAIR, & R. F. J. WITHERS

I: EXPERIMENTAL EVIDENCE ON THE RHEOLOGY OF COMPLEX ALLOYS AND ITS PHILOSOPHICAL SIGNIFICANCE

A. GRAHAM

I *Introduction*

COMPONENTS of machines, such as aircraft or gas turbines, may be made from complex alloys that contain perhaps ten elements, all of which affect their mechanical properties. In service use, the loading and temperature of the alloys vary in a rather arbitrary manner, and the many problems of designing to avoid failure in service are tackled by the industry mainly by the methods of direct development. These obviously lead to successful results, but they are expensive and laborious.

In a government research laboratory, individuals are less closely in contact with development experience than in industry and, if concerned with the problem, are at once impelled and more free to take up background research into problems of mechanical properties that are common to the class of commercial alloys with which the industry as a whole is concerned. Such a research poses the problem of providing systematic descriptions of the rheological behaviour of materials that are regarded by both industry and the academic world as being too complex for systematic and exact description.

Although the view that rheological behaviour is complex derives legitimately from the fact that it is concerned with the interplay of at least four independent variables, namely, stress, strain, time, and temperature, mechanical properties are governed by overall averages of a very large number of microscopic effects, and averages are invariably simpler

* Symposium at the Conference of Rheology, Manchester, 6-8 April, 1960

than the myriad details. Moreover, certain regularities that appeared in the published results of the rather small group of engineers that have studied the mechanical properties as a subject in itself suggested that if behaviour were systematically studied in terms of these four variables it would be found to be essentially simple.¹ The chance that appearances were genuine has led the author and his colleagues to spend about ten years in unravelling a tangled skein of regularity and to reach and support the conclusion that mechanical behaviour is no less ordered than other and better-understood aspects of nature. Acknowledgement is especially due to the contribution of Mr K. F. A. Walles.

The method followed has been the converse of that which is customary. Instead of progressively narrowing down and idealising the field of interest in direct pursuit of exact ideas, a search has been made for trends over the widest practicable extension of the field. A general mathematical principle has been progressively circumscribed to express the trends. The regularities found are of a rather unexpected nature, and it is not perhaps surprising that the work has been labelled as 'merely empirical' and without significance on one assessment and as 'academic and artificial' on another. In the present paper the nature of the regularities and of the experimental support for them are illustrated, and a discussion is offered of the fallacies of such general assessments.

2 Results

Certain ideas in the macro-rheology of metals run very close to ideas that have been put forward and developed in the field of non-metals by Dr G. W. Scott Blair and his colleagues. The present work stems from this double root, and it supports and extends a number of his conclusions. The regularities centre about an equation

$$\epsilon = C\sigma^\beta t^\kappa$$

put forward by Nutting in 1921 to represent the creep strain ϵ of pitch at time t when loaded with a constant stress σ . The quantities β and κ are numerical constants and C depends upon the temperature.

In order to take account of the experimental fact that behaviour cannot in general be represented by an algebraic equation of this kind, for the strain depends not only upon the instantaneous stress but also upon the way the stress has previously varied with time, the Nutting

¹ A. Graham, 'Phenomenological Theories of Creep', *The Engineer*, 1952, **193**, 198, 234; 'The Phenomenological Method in Rheology', *Research*, 1953, **6**, 92

EXPERIMENTAL EVIDENCE

equation needs to be regarded¹ as an integrated form of a more general equation such as

$$\epsilon = \int_0^{t_1} (t_1 - t)^\kappa \frac{d\sigma^\beta}{dt} dt$$

in which t_1 is a particular value of t . The equation is thereby made invariant with respect to time. The equation is allied to the fractional form¹ that Scott Blair has advocated. It represents an addition of elementary 'creep processes' which originate individually, with each increment of (σ^β) , at the instant each such increment occurs. Each contributes thereafter a progressively changing increment of strain.

For particular 'paths of loading', with σ as a function of t , as for example in ideal creep testing with stress constant, the equation reduces to the Nutting equation with an appropriate constant factor. The experimental support for this result shows that the use of an algebraic equation is legitimate when the loading path is fixed.

If attention is now confined to a particular loading path, the strain is not in general represented by a single Nutting term, but by the sum of a series of Nutting terms, thus

$$\epsilon = C_1 \sigma^{\beta_1} t^{\kappa_1} + C_2 \sigma^{\beta_2} t^{\kappa_2} + \dots$$

The regularities of especial interest, which the detailed studies have rather fully confirmed,² is that the exponents κ and β take only particular values. When the stress is constant for example, the exponents κ take values only from a sequence based upon the powers of 3, namely

$$\dots \left(\frac{1}{9}\right), \frac{1}{3}, 1, 3, (9), \dots$$

and the ratios of κ to β take values only from a sequence based upon the powers of 2, namely,

$$1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \dots$$

This algebraic formula is a straightforward inference from experimental results. The general formula contains a sum of integrals like the right-hand member of the second equation with the appropriate fractional values of κ and κ/β .

These regularities are seldom manifest immediately in experimental data on account of its scatter, limited range, and generally poor sampling, and because the effects of more than one term are generally superimposed. In favourable circumstances one or two terms predominate over the remainder: the fractional exponents are then closely

¹ A. Graham, G. W. Scott Blair, Discussion of paper by H. McCallion and D. M. Davies, *Proc. Inst. Mechanical Engrs.* 1955, **169**, 1135, 1137

² A. Graham, and K. F. A. Wallis, 'Regularities in Creep and Hot Fatigue Data', *Aeronautical Research Council Current Papers*, 1958, C.P. 379, 380

displayed. In other circumstances, their validity must be established by detailed analyses joined to an assessment of significance in terms of the confidence limits of the data.¹

The crucial test¹ has been to show that the laws lead to successful extrapolations of creep data over considerable ranges in time, for different sets of data grouped together to provide significant samples.

It appears from this work that the complexities of mechanical properties have less to do with the complexities of composition than with the presence of four variables and with the considerable range of these variables over which the pattern of regularity is extended. The formula can best be regarded as offering a rather high ratio of experimental facts co-ordinated to hypotheses made; the two series of standard exponents however appear to be, in some sense, fundamental.

3 'Fundamental' versus 'Empirical'

These conclusions are opposed to the view that it is almost inconceivable that all the complex phenomena concerned can have any simple explanation; that the phenomenological method that has been followed is mere curve-fitting without significance; that merely to represent results by a formula does not meet the essential need of explanation; that work on commercial alloys is necessarily empirical; that progress can only be made by fundamental studies; that only work on pure metals and simple alloys can be fundamental; that a fundamental result must be based upon atomic factors; that fractional power laws are merely empirical; that only integral powers and exponentials can be fundamental; and so on. The list is probably long enough for illustration.

Most of these problems centre on the significance of the term 'fundamental'; and they appear to involve old controversies that have been well fought over in the past. The term 'fundamental' has clearly both a scientific and a metaphysical connotation. Scientifically, a fundamental result or principle is one that is common to a wide range of experimental conditions and on the basis of which a wide range of results may be predicted. A metaphysical connotation is involved when the common principle is regarded as a fact of nature and as an absolute truth.

If one adopts the scientific viewpoint, then an important consideration is that whether a result is empirical or fundamental depends

¹ Work to be published

EXPERIMENTAL EVIDENCE

entirely upon the point of view. An experimental uniformity of behaviour such as the present can be regarded either as an empirical target for prediction on the basis of hypotheses made on the atomic scale of magnitudes, or as offering itself the fundamental hypotheses for some field theory of the macroscopic scale of magnitudes. Those who are primarily interested in atomic phenomena are thus liable to make the opposite assessment to those who, like the author, are concerned with behaviour on the scale of engineering. A great deal of confusion can arise from the fact that the two groups of research workers stand back to back, on a common result, and face in opposite directions.

Another source of confusion can arise from the fact that any science has two distinct phases with an essentially inverse relationship between them, namely, a formative phase and a development phase. The formative phase consists of inductive, analytical, and often disjointed detective work, in which one possibility is tried after another in a search for unifying principles and basic elements. It is not strictly logical in a classical sense and involves working inwards, from an imagined generality spread over a range of personal experience, along the lines of greatest probability, towards a centre.

When success is achieved, however, there is a complete change of procedure. The stage of formal development then follows in which the principles and elements are taken *by definition* as fundamental, and the science then develops outward and backward along the general course by which it came.

A great deal of misunderstanding can arise between those who are working outwards from familiar basic principles they accept without question and those who work inwards also, and are prepared to entertain some adaptation of basic principles. The necessity for such an adaptation in rheology is clear from the fact that, as Scott Blair has pointed out, some rheological properties are not physical properties in the ordinary sense of the term, for they are altered by the process of measurement.

Another source of difficulty is the canonisation of hypotheses: a promotion of respected hypotheses that have successfully co-ordinated a limited range of experience to the status of fundamental truths that must be preserved inviolate throughout all reasoning. In the present context, the hypotheses that appear to have been canonised are those that give rise to exponential laws and integral power laws. Canonisation evidently arises in the transference of ideas from master to pupil.

It inhibits thought because, in comparison, every other consideration is made trivial. Experience shows of course that the most revered hypothesis of one age and field of study turns out later to be only a limited part, or special case, of a wider concept.

A tendency to canonise is particularly associated with the concept of purity. The view exists, although fortunately it is dying, that work on pure materials is more scientific—more valid—than work on everyday materials. It is evidently a vestigial survival from the time when those who would now be called scientists studied the heavens, which were believed to be perfect and unchangeable, and tried to relate the best on this earth to the infinite. The mud, cheese, and metal, of present day rheology were earth-bound and profane. The present results cast doubt upon whether purity as such has any rheological significance.

4 A philosophy of the Phenomenological Method

The philosophy of the early days of quantum mechanics, of the Jeans-Schrodinger-Eddington-Stebbing period, appears to be a valuable guide to the labour of developing rheology from a craft into a science. One of the lessons that was learned with great difficulty, but all the more firmly in that period was the futility of attempting to visualise, in any final detail, the ultimate structure of matter in terms of the familiar pushes and pulls, billiard balls, and so called realities of everyday experience. Atoms, electrons, and the like were derived concepts, and had the status of artefacts which were not to be accorded any real existence in their own right. The business of exact science was to deal in observables alone—with the numerical results of experiment—and to specify numerically the experimental relationships between observables. A scientific theory consisted of mathematical symbols, operational rules for transforming one symbol into another, and a set of rules for interpreting results in terms of observables. Physical pictures and models were either scaffolding or metaphysics.

Whether this is the only valid workaday philosophy of science is a matter for debate, but at least it enabled atomic theory to escape from the trammels of its anthropomorphic origins. It is also the philosophy of the phenomenological method that has led to the present results. The philosophy is especially rejected by those who are fascinated by the various physical events that may be seen to happen before their very eyes; but the visual-model approach appears to belong to the early formative period of enquiry. The quantum-mechanical philosophy appears to be

EXPERIMENTAL EVIDENCE

adopted without difficulty in the later development period. Those who describe the rheology of metals in terms of patterns of dislocations, for example, readily transfer to the highly-formal theory of elasticity and Newton's laws of motion. The work here in review shows that progress can be made by a direct search for the mathematical expressions that adequately represent a class of experimental graph, and for the mathematical operations that enable passage to be made from one experimental graph to another.

To attempt to rank this procedure on a fundamental versus empirical evaluation is to obliterate all discussion with the metaphysical positive and negative infinities of a 'fundamental truth' and 'merely empirical device'; and it is quite unscientific.

5 *An application of the Phenomenological Method*

The view that research is a study of relationships rather than a pursuit of absolute truth permits a greater mobility of speculation. For example, the second equation

$$\epsilon = \int_0^{t_1} (t_1 - t)^{\kappa} \frac{d\sigma^{\beta}}{dt} dt$$

is a particular case of the general principle of linear superposition according to which any behaviour or function may be regarded as being composed of a distributed sum of repetitions of a unit behaviour or function treated as indivisible entities on a smaller scale of magnitudes. The 'unit function', t^{κ} in the above expression, may be any physical function whatever, and the 'unit cause' $(d\sigma^{\beta}/dt) dt$ may be any physical cause. The situation in nature to which the equation corresponds appears at first sight to be more akin to biological growth than to inanimate physical change—more akin to the growth of an organism by a progressive aggregation of parts with each part itself growing and changing in a uniform and characteristic manner. The distinction between biological growth and inanimate physical change is unlikely to be a genuine one, however, for rheology and biology are both concerned with regroupings and rearrangements of the same kinds of atoms. The biological parallel illustrates in passing the common misconception that worthwhile and reliable theories can only be built up from consideration of the behaviour of atoms.

The biological parallel is by no means fanciful or irrelevant. While discussing with Dr Scott Blair the mathematical bases of the power law, he drew my attention to an argument in a paper by von Bertalanffy

which appears to have originated with Julian Huxley. Huxley was interested in the ratio of sizes of two parts of the same growing animal; and he showed that if two independent quantities each growing exponentially are compared, then the ratio of their magnitudes at any instant follows a power relation. He supported the biological power law of relative growth by many examples from different animals. A simple adaptation of his argument leads to a similar sequence of exponents to those encountered, as previously mentioned, in the study of complex alloys.

In the comparisons of body size that led him to the power law, Huxley was concerned with the relative size of one finite part of an animal, say a claw, with the size of another finite part, say the shell, each being regarded as a uniformly-growing whole. He then discusses the case in which there is a progressive change of growth distributed over the animal, and quotes D'Arcy Thompson's observation that the forms of different species of fish or of different species of crab are geometrically related to one another by a regular Cartesian transformation. It appears, as Huxley remarks, that the different species have diverged in their evolution by the differing development of growth tendencies centred in different parts of the animal.

The use that Huxley and his successors have made of this reasoning may have a special rheological significance. The results offer a mathematical analogue, and perhaps even an actual equivalence in a chemical sense, between the deformation of an inanimate body in a mechanical environment and the changing form of a developing organism in a biological environment.

Ministry of Aviation
National Gas Turbine Establishment
Farnborough

II: NEW OUTLOOKS IN RHEOLOGY *

G. W. SCOTT BLAIR

'There is no escape from philosophy. The question is only whether a philosophy is conscious or not, whether it is good or bad, muddled or

* I am indebted to Mr A. Graham, Mr R. F. J. Withers, Professor M. Reiner and Professor A. Katchalsky for helpful discussions; and to Dr P. White for information about the history of the definition of fractional differential coefficients.

NEW OUTLOOKS IN RHEOLOGY

clear. Anyone who rejects philosophy is unconsciously practising a philosophy.'¹

INDEED it seems that those physicists who most hotly deny having any philosophy are inclined to be the first to condemn new ideas which contravene their own metaphysical presuppositions and such prejudices serve as hindrances to the development of the younger branches of physics.

Although the mechanistic beliefs of the nineteenth century are no longer accepted consciously, many physicists have not yet fully grasped the implications of the change in outlook. As Bridgman² has put it:

Many will discover in themselves a longing for mechanical explanation which has all the tenacity of original sin . . . Just as the old monks struggled to subdue the flesh, the physicist struggles to subdue this sometimes nearly irresistible, but perfectly unjustifiable desire.

The new outlook has been well expressed by Johnson³:

If symbols however indeterminate can be fitted together in functional form and in equations predicting what actually happens, we regard with tolerant but superior agnosticism our ancestors' desire to label those symbols and attach them to pictures of 'things' behaving as little miniatures of large-scale characters or bodies.

But many physicists have hardly yet reached this stage.

We have no precise definition for science as a whole, but for a description of the scope of intelligent scientific activity, one might well combine two well-known definitions, the one of intelligence and the other of science. Science is concerned with perceiving relations between those elements of our experience which are common or potentially common to all normal people.

Rheology is a branch of physics which deals with the deformation and flow of matter. As in other branches of physical science, there are four principal kinds of relation which we try to establish.

(a) *Relations between the behaviour of gross matter and that of its molecular or atomic constituents.* Rubber is believed to contract elastically because of its long chain molecules, pivoted to rotate at a certain angle, and buffeted about by neighbouring molecules into chance configurations. This model enables us to calculate the elastic behaviour of rubber at various temperatures. In a sense, this model is like the system of wheels

¹ K. Jaspers, *The Way to Wisdom*, London, 1951

² P. W. Bridgman, *The Logic of Modern Physics*, New York, 1928

³ M. Johnson, *Time, Knowledge and the Nebulae*, London, 1945, p. 160

in Mr Withers's black box (see p. 282), but in this case (a), we really believe that there are wheels, just as we believe that there are invisible genes which form parts of the visible chromosomes. We hope to get more and more evidence for their real existence as science progresses.

The rheological behaviour of gross matter usually depends on its molecular and atomic structure, so this type of relation is most valuable. In rheology, however, it is not always possible to find such relations, nor is it right to assume that all aspects of macrorheological behaviour are always entirely dependent on molecular or atomic structure. As an extreme case, two samples of sand, chemically identical, will behave quite differently when subjected to the same stresses if their particle size distributions differ appreciably. The fineness of grinding of clays much affects their plasticity. Here the causes may be partly at a molecular level, though certainly not entirely so.

(b) *Relationships between the behaviour of gross matter and that of other simple systems (models) which follow approximately the same equations, where no causal link exists.* We can observe light only when it is associated with matter and, in certain experiments, light then behaves like a wave. In certain experiments milk curd behaves like a particular system of viscous and elastic elements called, by rheologists, 'a Burgers Body' and shown in Figure 1 (see Scott Blair and Burnett¹).

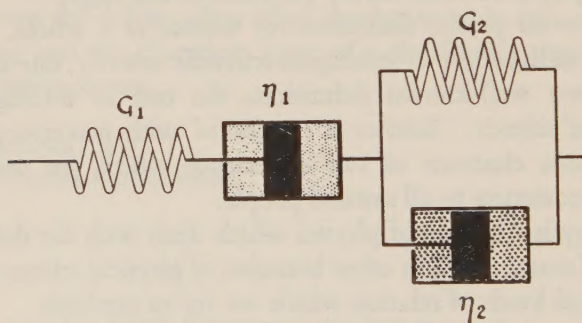


FIG. 1

An electric clock behaves like my watch, with all its wheels, but we do not expect the find to same mechanism within the clock. A failure in a distant power station will stop the clock but will not affect my watch. There are relations between light and waves; curd and Burgers Bodies; and the pointers on the black box, and the wheels in my watch.

¹ G. W. Scott Blair, and J. Burnett, *Brit. J. Appl. Phys.*, 1959, **10**, 15

NEW OUTLOOKS IN RHEOLOGY

(c) *Relationships between phenomena which at first sight appear quite unconnected*, but which are, in fact, linked, although no 'cause' is known: e.g. the forces which keep planets in their orbits and those that cause apples to fall to the earth. Such relationships are not common in rheology.

(d) *Relationships between the behaviour of matter under varying intensities of the same or comparable physical stimuli*. Under certain conditions, if we increase the stress on a milk gel, the strain which recovers when the stress is removed will increase proportionally.¹ Hence, as Mr Withers points out, we can calculate a useful rigidity modulus from many different observations. Under other conditions, the rate of strain is proportional to stress and a viscosity is a useful and meaningful entity.

Some rheologists dogmatically believe that one or other of these, generally (a) and (b), are the only 'sound' types of relation in our Science.² Enormously valuable as their contributions have been, this is just not true. Few rheologists stick entirely to type (a)—there are far too many phenomena for which an explanation cannot be offered in molecular terms. Type (b) explanations are accepted so long as the 'wheels' consist of ideally elastic (Hookean) springs, perfectly viscous (Newtonian) dashpots³ and (St Venant) sliders, representing static friction. Nor is one supposed to believe that these 'parts' really do exist. This is similar to the situation with respect to some of the complicated machinery of pre-Copernican cosmology. As with epicycles, so with dashpots and springs: if we postulate enough of them, we can always account for any experimental situation.

Spring and dashpot models lead to exponential equations, which, in the simplest cases, means that stresses will relax at constant strain and strains will build up under constant stress at rates proportional to themselves at any moment of time. The rates do not depend on the earlier stages of the experiment, since the moduli and viscosities remain constant. If, in fact, in more complex materials, the process of stressing or straining is altering the values of these properties, the simplest assumption would surely be to suppose that, in straining under constant stress,

¹ Ibid.

² Two American authors have recently shown much concern that they are not able completely to describe the behaviour of sand-water mixtures in terms of atomic arrangements. See W. A. Weil and W. C. Ormsby in *Rheology: Theory and Applications III*, Ed. F. R. Eirich, New York 1960, Ch. 8.

³ A 'dashpot' is a damping device consisting of a piston operating in a cylinder of viscous oil.

for example, the rate of strain, $d\epsilon/dt$,¹ would be proportional not to the strain at time t but to the average rate of strain since the time when the straining started, or

$$\frac{d\epsilon}{dt} \propto \frac{\epsilon}{t}$$

and this is a power equation. Writing σ for the constant stress and k for a fractional exponent (normally $1 > k > 0$) we have:

$$\epsilon = A\sigma t^k$$

and this equation does not give invariant magnitudes for simple properties like viscosities and elastic moduli. Such a 'model' cannot be precisely expressed in terms of springs with Hookean elastic moduli and dashpots with Newtonian viscosities. For a Hookean spring, $\sigma \propto \epsilon$, and

for a dashpot, $\sigma \propto \frac{d\epsilon}{dt}$. For the many complex materials which follow

power equations, we can write $\sigma \propto \frac{d^\mu \epsilon}{dt^\mu}$ where $1 > \mu > 0$.

Let us compare this type of treatment with the classical exponential treatment. The simplest general model to consider is that of Burgers, as shown in Figure 1. We may consider either the relaxation of stress at constant strain, or retardation of straining under constant stress.² Let us consider the former and plot the viscosities of the dashpots in the 'series' part of the model (η_1) against the moduli (G_1) of the springs.

Consider any one sample, X, of a set X, Y, Z. (Fig. 2a). This is placed so that the cartesian co-ordinates represent its shear modulus (G) and its viscosity (η) respectively. If we join X to the origin, then $\tan \theta$ will represent its relaxation time. (We can now leave out the suffixes, since the argument applies equally well to retardation times.)

Now consider a set A, B, C, of much more complex materials, whose behaviour cannot be easily represented by dashpot-spring models. Figure 2b, though superficially similar to 2a, is really very different. Again, one axis represents $(\text{stress}, \sigma)/(\text{rate of shear}, d\epsilon/dt)$ and the other $(\text{stress}, \sigma)/(\text{amount of shear}, \epsilon)$, but here the co-ordinates are polar. Viscous liquids may be shown as points along the vertical axis and elastic

¹ For the sake of simplicity, the same symbols are used for tensile and shear stresses and likewise for strains. These also conform to the usage in Mr. Graham's paper. This should be easier for the non-rheologist reader, though it is not to be recommended for strictly rheological papers.

² The ratio η_1/G_1 is called the 'relaxation time' that of η_2/G_2 , the 'retardation time'.

NEW OUTLOOKS IN RHEOLOGY

solids as points along the horizontal axis, at distances from the origin which represent the magnitudes of η and G respectively. The points A, B, C, which represent complex systems, are placed in such a way that the angle ω made by the line joining them to the origin represents the

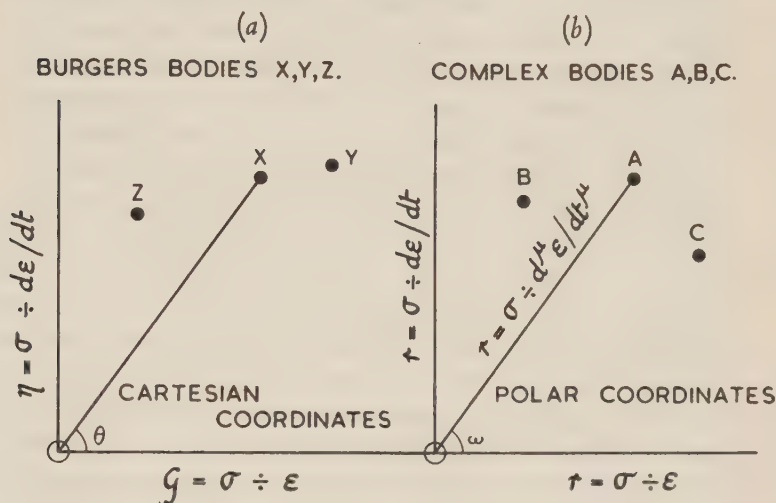


FIG. 2

position of the sample in a continuum lying between the viscous condition and the elastic condition. The radius vector OA is analogous both to η and to G , for if we represent it by χ we can write

For viscosity: $\eta = \sigma \div d^1\epsilon/dt^1$ or $\sigma \div d\epsilon/dt$

for elasticity: $G = \sigma \div d^0\epsilon/dt^0$ or $\sigma \div \epsilon$

for complex

systems: $\chi = \sigma \div d^\mu\epsilon/dt^\mu$ where $1 > \mu > 0$.

It is clear that the value of μ will depend on $\sin \omega$, varying from zero for an elastic solid to unity for a viscous fluid. Thus χ is intermediate between η and G , in the sense that the differentiation will be done μ times, where $1 > \mu > 0$. Furthermore, the length of OA, which represents the second polar co-ordinate, will represent the intensity of the property, χ , which approximates to viscosity when $\sin \omega \approx 90^\circ$ and to an elastic modulus when $\sin \omega \approx 0^\circ$. It will therefore have dimensions $ML^{-1}T^{-(2-k)}$, such that $k = 1$ for liquids and $k = 0$ for solids, and has intermediate values for the complex materials.

Now it is clear that, if we wish to compare, say A and B, we cannot just say $\chi_A > \chi_B$ as we might say $\eta_A > \eta_B$, or $G_A > G_B$. This would constitute the error in dimensions of which we are sometimes accused,

but there would seem to be no record of any rheologist making such a mistake. We can, of course, compare directly any set of samples having the same value of k (say 0.2, which happens to be a common value for high polymers). In other cases, we must always compare samples in terms of both k and χ taken together, i.e. of the position of the points on the diagram. This is a novel idea, but in no way an unsound one. Such samples are too complex in their behaviour to be compared along a single scale. Either they must be defined by a very large and perhaps infinite number of 'physical properties', or they may be compared in terms of their relative positions in a diagram such as that shown in Fig. 2b. If strain and stress are not linearly related so that a stress exponent is also needed, they are defined by means of their position within solid figures.¹

This method of treating certain data has two principal advantages. First, in many cases, as Scott Blair and others have shown^{2,3}, in order to obtain comparable goodness of fit, considerably fewer parameters are needed than if a series of exponential terms is used. Secondly, a much closer link is established in this way, with psycho-physical studies of the assessment of rheological properties of industrial materials by handling them. Such subjective assessments will long continue to be of importance in many industries. An earlier account of these ideas was given in a paper in this journal some years ago.⁴ Criticisms on dimensional grounds were ably refuted, at about the same time, by Dingle.⁵

A criticism of such methods more serious than that of dimensional inhomogeneity, concerns the definition of differential coefficients of fractional order. A concept of this kind must be precisely definable. One might start, for example, by trying to define the operation $d^{\frac{1}{2}}\gamma/dt^{\frac{1}{2}}$ such that, when carried out three times, the result is the same as a single differentiation. But we must then show that the result of applying such an operation is unique, or at least leads to only a limited number of alternatives; and this is not the case. It is better to follow Riemann's integral form of definition (inherent in Mr Graham's treatment) which gives a perfectly consistent and intelligible result for any real fractional order, rational or irrational.

¹ G. W. Scott Blair and B. C. Veinoglou, *J. Sci. Instrum.*, 1944, **21**, 149

² G. W. Scott Blair, B. C. Veinoglou, and J. E. Caffyn, *Proc. Roy. Soc. (A)*, 1947, **189**, 69

³ G. W. Scott Blair, and J. E. Caffyn, *Phil. Mag.*, 1949, **40**, 80

⁴ G. W. Scott Blair, *this Journal*, 1950, **1**, 230

⁵ H. Dingle, *Phil. Mag.*, 1949, **40**, 94

Although it is harder to envisage differentiating, say, $1/\pi$ times than to imagine a half, or a third of a differentiation, one must assume that the space in Figure 2(b) forms a continuum. (This is a point which was raised by Dedekind.) If we suppose a line, corresponding to a set of samples having the same distances from the origin starting at $\sin \omega = 0$ (elastic solid) and gradually rotating towards $\sin \omega = 1$ (liquid), this line will pass continuously through all intermediate angles, not only those corresponding to rational fractions. It is unthinkable that there will be some kind of 'exclusion principle' forbidding specified angles.

The idea of comparing the behaviour of complex bodies in terms not of distances along a line (a scale), but by their positions in a space of two or more dimensions, is worth further consideration. But it is clear that such treatments will not lead us to simple models of rheological behaviour!

It has been objected that the quantities here proposed for the specification of the rheological behaviour of certain materials are not tensorial and, therefore, do not have the invariance required of fundamental physical properties. This, of course, depends on what is meant by 'fundamental physical properties'. As a concession, I have felt it best to describe such entities as 'quasi-properties'; though I am not sure whether this concession was really a wise one. There are many situations in physics for which the tensorial type of invariance is highly advantageous, but I cannot see that establishing relations between the rheological behaviour of complex materials is one of them. As an illustration, some elastic materials do not obey Hooke's law over any appreciable range of straining, but a power relation is found to hold between stress and strain (Bach's equation). It is clear that this power equation will not be valid for very small strains, when Hooke's law is probably obeyed. But this in no way precludes the use of Bach's equation for all but the smallest strains. Another case is the fairly wide use now being made of Casson's equation¹ which gives a linear relation between square-root of stress and square root of shear-rate in material flowing through a capillary, or between concentric cylinders. This does not allow of a reversal in the direction of the flow.

I would conclude with a word about Mr Withers's type (1) and type (2) explanations (see p. 283). I take it that what I have called (a) and (b) relations offer type (2) explanations, whereas (c) and (d) are of type (1). Sometimes the question whether a model is supposed to represent real

¹ N. Casson, in *The Rheology of Disperse Systems*, Ed. C. C. Mill, London 1959, Ch. 5

components of a system, is left open; but, as time goes on, the tendency is more and more against reification of such models.

Mr Withers asks if there are any hopes of finding some explanation of type (2) class, i.e. perceiving relations of my types (*a*) or (*b*) with power laws and fractional differential coefficients. He agrees with me that this is not *essential* and I agree with him that type (2) explanations have proved so valuable in the past that it would be well worth trying.

I feel that Mr Graham's contribution comes nearest to making such a link. I feel pretty sure that if we ever succeed in finding an explanation of the type (2) class for the power equations, it will come through the simple mathematical fact that a power equation results from the combination of two exponentials; but I cannot yet see just how this is to be done.

National Institute for Research in Dairying
University of Reading

III: EXPLANATIONS AND MODELS IN SCIENTIFIC THEORY CONSTRUCTION

R. F. J. WITHERS

I *Introduction*

THE work of Graham and Scott Blair has been criticised in the past on various grounds. Graham has enumerated some of these objections. I want to concentrate on a few of them:

- (1) That it is inconceivable that all complex phenomena can have a simple explanation.
- (2) That the phenomenological method that has been followed is mere curve-fitting without significance.
- (3) That merely to represent results by a formula does not meet the essential need of explanation.
- (4) That progress can only be made by fundamental studies.
- (5) That a fundamental result must be based upon atomic factors.

These criticisms suggest to me that a discussion of the rôle of generalisation and explanation in science might be useful with respect to their work and this will also involve a discussion on the rôle of models in scientific theory construction.

SCIENTIFIC THEORY CONSTRUCTION

My impression is that whenever a controversy over scientific statements arises a close examination of the type of statement being discussed and a study of the statement's exact position in the ordered hierarchy of scientific statements focuses the discussion on to the way that that statement is functioning within the scientific theory. To do this is always a profitable pursuit and is one of the main functions of the philosopher of science.

2 *Sketch of Statements in a Rheological Theory*

In order to arrive at some idea of the rôle played by the extension of the Nutting equation in rheology I shall consider a statement of its function as expressed by Scott Blair and Burnett in a paper¹ they published in 1959. I shall classify the statements made in that paper under the scheme described by Woodger.²

Scott Blair and Burnett reported *observation records* about what happens when small pressures are applied to one end of a column of renneted milk ('curd') in a U-tube and the displacements are magnified a hundred fold by means of an alcohol index in a horizontal capillary. At the first stage of the discussion no theoretical terms seem to enter. However, soon the authors calculated the overall shear modulus G . This we are told is calculated from the movement of the alcohol index of recovery h , the density of the manometer liquid ρ , the radius of the U-tube R ; the length of the column of gel L and the radius of the recorder capillary a , from the equation

$$G = \frac{R^4 \rho g}{8La^2} \cdot \frac{P}{h}$$

The overall G can be calculated from an experiment in which the sample is loaded for, say, one minute then allowed to recover for, say, two minutes. The whole of the recovery, almost complete in two minutes, gives us h .

In the second type of experiment readings are taken at frequent intervals during the two minutes of recovery and so they calculate a series of unrecovered displacements. This enables G to be divided into two parts, immediate and slow.

Now it is obvious that more is involved in this procedure than mere observation. The equation from which G has been calculated must

¹ G. W. Scott Blair and J. Burnett, 'On the Creep, Recovery, Relaxation and Elastic "Memory" of some Renneted Milk Gels', *Brit. J. App. Phys.*, 1959, **10**, 15-20

² J. H. Woodger, *Biology and Language*, Cambridge, 1952, Chs. 1 and 2

represent some other type of statement in which G is a definition for a complex relationship. Certain of the terms in the equation are also definitions of other relationships. My remarks about G will be applicable to these terms.

The definition of G is based on a number of observations which suggest a general relationship, a generalisation of observation records, expressed by the right-hand side of the equation which is believed to exist in a repeatable fashion. The relationship is sufficiently important for it to be abbreviated to the single letter G and to be known as the shear modulus.

The authors go on to describe the result of calculating four parameters: the immediate modulus, the delayed modulus, the internal damping viscosity and the external viscosity. They notice that these parameters 'behave approximately as a Burgers body'. It seems to me that the impetus to discovering this behaviour must have come from the use of a model, that of springs and dashpots, suggesting the making of observations about the four parameters that were calculated. The results obtained were the test of the spring and dashpot model, where this was functioning as an *explanatory hypothesis* suggesting the making of fresh observations. Let us, therefore, consider what is involved in inventing an explanatory hypothesis.

If we consider the sort of statements that form explanatory hypotheses we can see that there are two main types. Let us consider a simple example of a situation requiring explanation. We have a black box with a dial on the front having numerals at various points and two hands which rotate at different speeds. Can we devise an explanation for the behaviour of the hands and in particular can we predict when the hands will coincide and can we perhaps give an estimate of the times on our watches when the coinciding will occur? A scientist could look for an explanation of the behaviour of the hands in two ways.

(i) He could calculate the angular velocity of both hands and work out an expression relating the two in such a way that the required prediction can be made. This expression would constitute in one sense an explanation of the behaviour of the hands (I am going to call it a *Type 1 explanation*). This explanation could tell one a lot about the movements of the two hands and could predict various relationships between them. The question I think that might be considered is whether the extension of the Nutting Equation is an example of a *Type 1 explanation*.

SCIENTIFIC THEORY CONSTRUCTION

(ii) Or, a scientist could first remember that he has seen other systems, such as the one he has in the watch on his wrist which he knows contains cogs, balance wheels, springs, and so forth. Using this as a conceptual model he could then invent statements about the presence of such entities inside the black box, and their relationships to one another and their relationships, causally, to the movement of the hands. From these statements he could predict the behaviour of the hands as before. He could also predict, for example, the existence of cogs, etc., as the parts of the black box system. He could make predictions about the number of teeth on the cogs which he suggests exist in the black box and he could test these predictions by opening the black box. Notice what he has done. He has used his own watch as a model system and the statements about the parts of his watch he has transferred to the black box system. This may be successful in predicting the behaviour of the hands but it may be falsified if he finds no cogs inside the box. It may, for example, turn out that the black box is run by electricity using an alternating current. But I suggest that until he has tested his theory his statements about the hypothetical cogs will give him reasonable explanation of the behaviour of the hands on the front of the black box. In other words—it is the same as a *Type 1 explanation*. He can only test the complete adequacy of the model in terms of the predictions of the cogs and other parts.

Another thing to notice is that this type of explanation involves inventing the existence of parts of the black box and their hypothetical relationships and it is because of the invention of these parts that totally *new* observations can be made. Let us call this type of explanatory hypothesis a *Type 2 explanation*. This is a convenient point to indicate that there is a fourth type of statement found in science, *consequences of explanatory hypothesis*, i.e. statements about the hypothetical cogs.

I do not want to discuss the merits of either type¹ of explanation, but is it only fair to say that most scientific theories progress along the lines of *Type 2 explanations*. The behaviour of chemicals in a test tube is explained by the statements about molecules. These can be explained in terms of statements about atoms and valency and these can be explained in terms of statements about electrons and protons, so that one arrives, in a well developed science, at a theory with several layers of *Type 2 explanatory hypotheses* from which an original observation

¹ A somewhat similar analysis of explanatory statements has been made by B. Ellis, 'A Comparison of Process and Non-process Theories in the Physical Sciences', this *Journal*, 1957, 8, 45-56.

record is derivable and from which new types of observation records can be made.

3 *The Use of Models in Science*¹

In the example I have just given we used the model of our own wrist watch for one reason. This was because we have a description of its behaviour ready to hand as a calculus. The calculus consists of statements which have been tested in the past and found satisfactory for our watches. The calculus itself will consist of statements expressing relationships between variables and in order to derive consequences from it which can be tested those variables must be replaced by subject matter signs. In other words the calculus must be given a semantical interpretation *at least* and only *at least* at the lower levels of the calculus. It must have a semantical interpretation at those levels where we wish to test it. However, when we use this calculus as a model the first problem that crops up is whether the semantical interpretation, appropriate to the sciences of wrist watches, is suitable for the new situation. In the case of the black box being an electric clock it will soon turn out that the interpretation is falsified. Note that it is the interpretation which is incorrect—not the calculus—for that enables us to predict. We are then faced with the choice of looking for a new interpretation of the variables in the calculus or trying to think up a new calculus with its own interpretation.

In any case, the important point is that the model was used in the first place because it had its own calculus, and it was a system of ready to hand statements that we were looking for. Provided we can invent a calculus we do not *need* to look for a model. In practice it is difficult to invent calculi, and the heuristic desirability of models for supplying calculi has unfortunately changed the emphasis from realising the importance of the calculus to believing the importance of the model. The calculus is the most important thing for the development of a theory, the model—or rather the semantical interpretation of a model—may actually be misleading as in the billiard ball model for fundamental particles.

Now a danger in the development of a science is that a good science may have a well established series of layers of explanatory hypotheses

¹ A more detailed account of the use of models in science is to be found in E. H. Hutten, 'The Role of Models in Physics', *This Journal*, 1954, 4, 284-301, and M. B. Hesse, 'Models in Physics', *this Journal*, 1954, 4, 198-214.

SCIENTIFIC THEORY CONSTRUCTION

of *Type 2*. Because they are established they fit most fields and have been successful in the past. If a new field of observation records arises the generalisations of them may be arranged as the *consequences* of using as a model calculus an established theory which has been successful and is interpreted in another field. Provided that there are no counter-examples, progress proceeds by taking over the old theory as a model into a new field and looking for entities in the new field which behave like the parts at the various levels in the old field. In other words, we attempt a new interpretation of the calculus by looking for new subject matter signs to replace the variables in the model calculus.

If counter-instances are produced or if predictions using the old theory do not quite fit the generalisations of observation records in the new theory an attempt can be made to modify the model. Generalisations of accurate observation records in the new field may give rise to discrepancies when tested if the old theory is used to interpret or derive them. Further, when one has a satisfactory system this can be sufficiently powerful for it to dictate the sort of generalisation that is made. If this is the case and counter-instances are still produced specific *ad hoc* hypotheses can be invented to cover the discrepancies from experience. In the end the calculus becomes extremely cluttered up with these *ad hoc* hypotheses.

Further, a situation may arise in which a scientist has attempted a *Type 1 explanation* which works with greater elegance and greater precision in prediction. In this case the problem becomes acute if the *Type 1 explanation* is not derivable from the older theory used as a model. The scientist is then faced with the problem of whether to use the *Type 1 explanation*, because of its efficiency, empirically, to abandon it, or whether to look round for another model which might help him to invent *Type 2 explanation* from which this particular *Type 1 explanation* is derivable. So far as I can see there is no reason in principle why this should not be attempted, other than the usefulness of the old model and the natural conservatism of scientists. If it were achieved it may involve the over-throw of a pre-existing theory. However, it is worthwhile remembering that it is only a model calculus that is being over-thrown and the process may possibly end up by being responsible for a more adequate theory. The new theory incidentally should have as one of its consequences some of the successful statements of the old model. This is the way in which real progress in science is achieved.

4 *Discussion of the Nutting Equation in Rheology*

Let us now examine the situation as seen in the Scott Blair and Burnett paper which reflects the present position. The authors say in the discussion that they have found it convenient to *describe* their experiments in terms of models of dashpots and springs. It is worth noting that by description here they mean that they have *explained* their results in terms of a *Type 2 explanation*. However, they note that there are limitations in this particular explanatory hypothesis. For example, they ask how the principle of superposition is to be reconciled with the power equation relating dissipating stress to time? The principle of superposition was derivable from the spring and dashpot model, but the power equation was obtained from experiment. It is true that this could have been derived from extending the spring-dashpot model to an infinite number of springs and dashpots related exponentially, but otherwise was not obtainable. The calculus would therefore become exceedingly cumbersome, and unless it was made cumbersome it would not lead to accurate predictions for relating dissipating stress with time.

In an attempt to offer an explanation which will cover the difficulties between linear superposition and the power equation for relating dissipating stress and time, the authors turn to their own solution. This is to derive the relationships from an extension of the Nutting Equation. Now this seems to me to be an attempt at a *Type 1 explanation*. It can be tested in terms of deriving consequences and seeing whether the fit of the data does not significantly differ from the results expected on the basis of the extended Nutting Formula. I take it that this has been done and there is no doubt that as a *Type 1 explanation* the Nutting formula is satisfactory. Moreover, it is more satisfactory because of its having fewer parameters than the equations derivable from even a spring-dashpot model.

The problem then is not whether to accept this *Type 1 explanation*, but to relate it to current and future rheological theory. Graham points out in his paper that this is the point in discussing what he calls the difference between fundamental (already existing) and empirical (construction of *Type 1 explanations*) views. In relating it to present rheological theory one can proceed in two directions.

- (a) To look for a *Type 2 explanation* of their formula.
- (b) To look for consequences of their formula which predict an occurrence which would not occur under the spring-dashpot

SCIENTIFIC THEORY CONSTRUCTION

theory. A test of this consequence would be the deciding factor between the two theories. Can this be done?

I take it that the major question in the minds of most critics of the use of the Nutting formula by Graham and Scott Blair is that at the moment there is no *Type 2 explanation* of it.

It is a possibility that no satisfactory *Type 2 explanation* is forthcoming. In this case a search for examples of the application of the equation might be useful. Already it appears to me that a wide range of application has been shown. A similar situation can be seen in a new field of study—General System Theory—discussed by von Bertalanffy,¹ in which a study is made of those systems in which predictions can be made on the basis of isomorphic laws. A study of the characteristics of the complex systems which are covered by such isomorphic laws may give us a clue as to *Type 2 explanations* which might be useful in the various fields. In this way the relationship between power law in rheology and the power equations involved in the phenomena of relative growth in biology² might be related. One such model, applicable to Scott Blair's material and biological phenomena of growth would be a model based on the kinetics of enzyme activity, for the biologist believes that the differences in relative growth in different organisms is related to genetic factors affecting rates of enzyme activity in different processes which affect growth.

5 Conclusion

We must therefore return to the objections which have been mentioned in the beginning.

- (1) That it is inconceivable that all complex phenomena can have a simple explanation. My answer would be that there is no reason why even complex phenomena could not be *described* in terms of a simple generalisation, provided one can describe what one *wants* to describe without introducing all the complexities inherent in the material. This minimum type of description is the basis of all generalisation. If we went to describe the weight of a lump of sugar, or its behaviour as a falling body, then we leave out its colour, its taste, its use etc., and yet can still make meaningful generalisations about it so far as it is to be

¹ L. von Bertalanffy, 'An Outline of General System Theory', this *Journal*, 1951, I, 134-72

² J. H. Huxley, 'Problems of Relative Growth', London, 1932.

described as a falling body. The same sort of difficulty is found in description of human behaviour which is undoubtedly very complex, but advances are made when some of this complexity is shed. Then there is no reason why a *Type 1 explanation* which appeared simple should not be used to derive this simple generalisation. A *Type 2 explanation* will undoubtedly increase the complexity of the explanation, in so far as new concepts are involved in this.

- (2) That the phenomenological method that has been followed is mere curve-fitting I do not think in *itself* is a bad thing. If the curves enable more accurate predictions to be made and are more economical of parameters it represents an advance because the curve represents a *Type 1 explanation*. Unless it can be shown that a *Type 2 explanation* cannot, in principle, be suggested from which the *Type 1 explanation* follows, the formula for the curve should be accepted.
- (3) That merely to represent results by a formula does not meet the essential need of explanation. This objection rests on an understanding of what is meant by the essential need of explanation. If by this we mean the need for a *Type 2 explanation* then we can attempt to think of the direction of the next step in theory construction which is to *find* such as the *Type 2 explanation*.
- (4) That progress can only be made by fundamental studies is quite definitely a false concept. There are two sorts of progress, (a) the search for generalisations and *Type 1 explanations*, and (b) the search for *Type 2 explanations*. Finding of better *Type 1 explanations* may require the revision of previously existing *Type 2 explanations* as in my clock example.
- (5) That a fundamental result must be based upon atomic factors. This is only true so far as one sort of progress is made by searching for *Type 2 explanations*.

Finally, in case we feel that *Type 1 explanations* are unsatisfactory at low levels (low in the sense that they are close to observation records) at high levels we have to use *Type 1 explanations* until we can see a way of talking about still higher levels. A scientific theory is a calculus and we need not ask first whether the term in it refers to anything. The prime question is whether the formulae expressed at the higher levels are useful in prediction.

Middlesex Hospital
London, W.1

DETERMINISM IN CLASSICAL PHYSICS ★

G. F. DEAR

IN 1951 Professor Popper published two papers¹ in which he argued that classical physics is indeterministic in perhaps no less fundamental a sense than is quantum physics. Popper's argument was in two stages, the first being a purely philosophical preliminary to the second, which was a technical exercise in physics and logic. Firstly, he substituted for 'the metaphysical doctrine of determinism' the more precise and testable notion of complete, though 'finite', scientific predictability. Secondly, he tried to show that even in a universe which is 'prima facie deterministic', that is in one which is completely describable in classical terms, there can exist certain physical systems which are unpredictable.

This paper is divided into two main parts. In the first part (§§1-3) two possible scientific explicata of Determinism are analysed and assessed: predictability is found wanting and scientific lawfulness is shown to be satisfactory and replete. In the second part (§4) Popper's technical thesis is examined and rejected.

Although the two parts of this paper correspond in content to the two stages I have distinguished in Professor Popper's argument, only the second part is directly contrary to what he was concerned to establish. In taking predictability as a testable rendering of the determinist thesis, Popper was not necessarily committed to its adequacy, nor was he necessarily preferring it to scientific lawfulness. Rather, the intention was to attempt to meet the Copenhagen school of quantum theorists on their own ground by adopting without argument a philosophical orientation similar to their own, and thus to make more profitable comparison of classical physics with quantum physics.² However, if the argument of the first part of this paper is correct and incomplete predictability is allowed to be not necessarily incompatible

★ Received 10. ii. 60. The author wishes to acknowledge the useful conversations he has had with Dr G. Buchdahl, Professor R. B. Braithwaite, and Professor K. R. Popper while preparing this paper.

¹ K. R. Popper, 'Indeterminism in Quantum Physics and in Classical Physics', this *Journal*, 1950, I, 117-133, 173-195

² This is of course only my interpretation of what Popper intended.

with determinism, then even if Popper's technical thesis had been valid it would still have had no essential bearing on determinism. Predictability and determinism, it is here maintained, are different things.¹

I

Before proceeding to explicate determinism let us briefly consider, first, why determinism stands in need of explication, and second, what criteria an adequate explicatum must satisfy.

There are no laboratories, scientists or budgets devoted to determinism. And this is not the result of an oversight or of an unsympathetic D.S.I.R.: it will become clear below that whether the world is or is not deterministic logically cannot be the direct object of scientific enquiry. How then does determinism make contact with the contingent? Merely as some sort of higher-level scientific programme? No. Not merely. Though certainly it can be advanced as a programme. But it is worth noting that for a scientific (or meta-scientific) programme to be advanced at all it is necessary that the proposition which affords it rational direction² be rationally entertained, or rationally doubted, and hence that there exists some evidence, immediate or theoretical, bearing on it.

What kind of evidence has traditionally been cited in support of a determinist thesis? Not, roughly speaking, the small or large experimental evidences that bolster or topple specific hypothetical regularities and laws, but rather the fact of the existence of regularities or laws as such, without regard to their particular contents. It is by reflecting on the regularities as regularities that the idea of determinism is born: the idea that clocks, clouds, computers, human beings have something in common other than their patent materiality, vaguely, the lawful way in which they unfold in time. And as it is in the context of the *prima facie* conflict of this idea with the consciousness each man has of being a free person, rather than in the active pursuit of science itself, that the notion of determinism has developed much of its point and meaning the problem is this: to find an empirical thesis which is similar to the original determinist insight in range and content, which is sufficiently exact for the assessment of its truth or falsity to be in

¹ Cf. W. B. Gallie, *Free Will and Determinism Yet Again*, The Queens University of Belfast, 1957

² E.g. the programme 'find out whether all S and P' is directed either by the hypothesis 'All S are P' or by \sim 'All S are P'.

DETERMINISM IN CLASSICAL PHYSICS

principle possible, which is—for preference—simple, and which further, does not in rendering the original empirical, distort it in such a way as to gratuitously resolve the 'Free Will Problem'. Such transit from an unclear, prescientific insight to a decidable scientific hypothesis is not essentially and so necessarily 'psychological' in nature, the result of a random imaginative jitter or of a causal push: it *can* be rational, and the process is then usually referred to as 'explication'.¹

2

We take as starting point, or explicandum, the neutral :

By necessity are foreordained all things that were and are and are to come.²

Although Democritus intended this scientifically, as a bare proposition, it is of course not specifically scientific, or metaphysical, or, anything too particular. Speaking loosely, three main ideas can be distinguished in the explicandum: necessity, fore-ordination, and the universality of the reference of these two ideas, that is, their predication of *all* events. Whereas all theses that have traditionally been known as 'deterministic' have had the universal reference, the first two ideas need not both belong: theological determinism has usually placed fore-ordination or divine omniscience centrally, the necessity, if it is not rejected, being 'only' the necessity of a divine fiat; metaphysical determinism, as for example that of Spinoza, has dwelt on the necessity and retained the fore-ordination, if it has retained it, merely as a figure of speech.³

Roughly speaking, the scientific explicata we shall consider have emphases paralleling those above: predictability is empiricised fore-ordination, and scientific lawfulness maintains the element of necessity in the formal necessity or deducibility of later conditions given earlier ones, for all that only experience can tell which particular theories might in fact be 'laws'.

We shall begin by considering scientific lawfulness, which we shall refer to as 'determinism₁' (or d_1)⁴. A reasonably precise formulation is the following:⁵

¹ See R. Carnap, *The Logical Foundations of Probability*, Chicago, 1950, p. 5.

² Democritus (Plutarch, Strom. 7) ³ Leibniz, of course, realised both ideas.

⁴ For the original explicandum 'Determinism' will be used.

⁵ This is a somewhat modified version of that proposed by P. Feyerabend in a discussion in *Observation and Interpretation*, ed. S. Körner, Bristol, 1957, p. 182.

'The world is deterministic if, and only if, the class of all observables can be divided into sub-classes in such a way that for every sub-class there exists a theory which is deterministic with respect to a set of variables consisting of that sub-class plus, possibly, certain "hidden parameters", and which is true', where a deterministic theory is defined provisionally by:

'A theory is deterministic with respect to a set K of variables if, and only if, a conjunction C of statements to the effect that the variables of K have a certain value at the time t_2 can be derived from the theory together with another conjunction of statements, which again contains only the variables of K , and which asserts that those variables had some other value at some other time $t_1 < t_2$.'

Although d_1 is empirical, it is no ordinary scientific hypothesis. It is peculiar in that, being an assertion about the logical or mathematical form of true scientific hypotheses, it bears on the World only via these hypotheses. Its empirical content is that: the theories which truly describe the World are deterministic in form.¹ d_1 is itself descriptive, even though indirectly so.

What is to count as evidence for or against d_1 ? Clearly, we can conceive no stronger or more relevant evidence than the deterministic or indeterministic form of the presently accepted scientific theories. Of course, there can be no guarantee that these theories are 'true', in that fresh experimental evidence may always lead to their falsification. However, in so far as they are 'so far forth true' relative to all the evidence to date, a wider evidential base for a d_1 which is to be empirical cannot be conceived, since this base is as wide as science itself.

Determinism₁ is presently falsified if there exists one instance of an accepted theory which is not deterministic itself, and for which there are no grounds for asserting it to be reducible² to other deterministic theories. For investigators into determinism the accepted theories are to be found in the recognised journals and textbooks. Assessment of the truth of those theories, of the extent to which 'hidden parameters' are acceptable variables in a theory, and of the plausibility of attributing a distribution of errors about a predicted value of a variable to determinate, though perhaps uncontrollable, variations in the boundary conditions, are matters for scientists and not for

¹ Cf. L. Wittgenstein, *Tractatus Logico-Philosophicus*, London, 1922, 6, 32.

² E.g. in the sense of 'reduction' introduced by Kemeny and Oppenheim: 'On Reduction', *Phil. Studies*, 1956, 7, 6-19

DETERMINISM IN CLASSICAL PHYSICS

determinists₁. At present the evidence is against d_1 : quantum mechanics plus the Born probability interpretation have still to be reduced to a more basic deterministic theory.

In passing, it is worth noting briefly why determinism₁, although an 'all-and-some' proposition possessing both universal and existential quantification,¹ evades the 'metaphysical' label recently attached by J. W. N. Watkins to the superficially similar Law of Universal Causation.² Whereas it is, I suppose, still just possible for an empirically-minded metaphysician to insist that, all appearances notwithstanding, a cause will always be found for every effect if one looks long enough—though how and where to 'look' is of course not specified—yet d_1 cannot be so easily compromised. Theories are human artefacts. Their existence is not in doubt. Neither is their deterministic or indeterministic form. So that one need look only as far as the libraries.

For purely mathematical reasons we shall now loosen somewhat our definition of a deterministic theory (see p. 12). There are two difficulties: firstly, for the vast majority of problems the equations of mathematical physics can only be solved approximately,³ either by 'perturbing' a purely analytic solution of a simpler problem, or by solving numerically from scratch; secondly, if the time interval $t_2 - t_1$ is allowed to become infinite, mathematical singularities and paradoxes multiply rapidly. These difficulties are avoided by requiring of our theories not exact temporally-unbounded solutions, but rather solutions for a finite time interval obtained after a finite number of deductive steps: given precise values of the variables for the earlier time t_1 , it must be possible to derive, for a later time t_2 (where $t_2 - t_1$ is finite), values of the variables which are as close to the 'true' values as we please. Or, we require of a deterministic theory convergent solutions over a finite time interval.

But what if, for certain special problems for example, the solutions are not convergent?

Such difficulties are emphatically not for determinists₁, or indeed for pure mathematicians (once the non-convergence is recognised), but for scientists. There are two alternatives: either there is a proof of

¹ E.g. Roughly, d_1 can be written as 'All observation statements are subject to some deterministic theory'.

² See e.g. 'Confirmable and Influential Metaphysics', *Mind*, 1958, 67, 344.

³ Even when probabilities are not explicitly involved.

non-convergence, or there is not such a proof. If there is a proof, then the scientist must decide whether the non-convergence, or perhaps the complete lack of any sort of solution at all, is to be attributed to inadmissible idealisations in setting up the problem for the theory,¹ or whether rather, having a problem which is just not effectively covered by the theory (since it gives no solution), he ought not to look for a theory which *does* cover the problem. For the second alternative: if there isn't a proof of non-convergence and it is uncertain whether a solution exists, then either the scientist will look for a better theory, or he will 'rationally intuit' that a mathematical discovery will one day give him a solution. And as he knows the problems and has all the evidence, is it for us to quarrel with his judgment?

3

As an alternative explicatum, favoured by some, we consider now predictability, which will be referred to as determinism₂ (or d_2) and which initially is simply: 'All events are predictable.'²

Determinism₂ is sometimes advanced as though it were too obviously an adequate scientific explicatum of Determinism to require justification, especially if Determinism itself is conceived in theological terms, in terms of predestination or of omniscience. More often however, it is put forward either as being in some sense an analysis of what d_1 amounts to when direct reference to the—on this view—purely mediatory theories is avoided, or as being an empiricised version of a position which is otherwise 'metaphysical'.

That d_1 is content-full has been argued above. We shall now show that d_2 is significantly distinguishable from d_1 , and so is not acceptable as an 'analysis' of it, and shall briefly explore d_2 's inadequacy as an explicatum of Determinism.

At first sight, d_2 appears to possess an epistemological immediacy and a simplicity so conspicuously lacking, the reader may feel, in d_1 . For is not 'All events are predictable' of the familiar form 'All α are β ' and such that the instances ' α_1 is β ', etc., are virtually observation statements? Cannot it be very simply tested by placing a good scientist in any arbitrary observational field and noting his performance:

¹ And this, for example, is the case for three-body collisions in Newtonian mechanics: punctiform masses do not exist.

² Cf. the more exact 'finite' version of Professor Popper given in section 4 below. The analysis and criticism of d_2 which follows applies equally to it.

if he makes successful signals then *that* particular kind of α is β , that is, it is predictable? But d_2 is not at all simple.

An invariable first step in developing d_2 must be to understand by 'predictable': predictable by means of scientific theories. This is necessary if revolver-compulsion and/or successful Old Moore universes are to be excluded.¹ This first specification of the meaning of 'predictable' removes at once the epistemic simplicity of d_2 : d_2 is now a proposition operating on two different levels. On its own level—which we shall call the meta-level—it asserts the universal possibility of an 'intelligent' object, e.g. a scientist, successfully predicting any specified event. And the 'predicting' here is an activity performed in time and space by a contingent being, the success of a prediction being inferred from the intelligent object's behaviour. But this object is also a subject. At the lower level—lower because it is circumscribed, dependent on the experimental and other conditions obtaining at the meta-level—the subject or scientist performs logical operations within the appropriate scientific theories, and assesses the truth-value of his inferences against his observations.²

The fundamental objection to d_2 is that scientific theories play a part at both of these levels, and that because of the 'subjectivity' at the meta-level an infinite regress is generated. Any remotely acceptable d_2 is saturated through and through with theory at its own level, as well as at the lower level. Only thus can it be given the necessary universality and the anthropocentricity or 'subjectivity' inherent in d_2 be at least superficially mitigated. To secure this, there is always added to 'All events are predictable' the sweeping qualification 'in principle'. Any appeal to this qualification is an appeal to theory.

Firstly, let us suppose that our intelligent being is *very* intelligent, his grasp of scientific theory unrivalled, his observational acuity a by-word: what would we need to add to a finite number of successes to secure 'all α are predictable by a perfect scientist'? Spatial and temporal universality. If space-time is accredited—by the relevant theories—with homogeneity the inference is not very difficult. It corresponds to the early extension of the verificationist meaning

¹ It is worth noting here that since 'the scientific theories' are not specified, d_2 no less than d_1 , is an 'All-and-some' proposition and thus needs rescuing, just as much as does d_1 , from the 'metaphysical' label.

² The acceptability of 'errors' is assessed at this level also.

criterion to the past, and to distant points (e.g. Mars) which are unattainable for merely 'practical' reasons. On similar grounds it is recognised that *all* the events occurring during any one short period of time will never, unless the universe changes radically, be predictable, because the number of events will always be larger than the number of perfect scientists. In short, all those events are admitted which could be predicted were a scientist or an observer-calculator to be present, and for which the presence of such an observer-calculator at a spatio-temporal point appropriate to observation and prediction of the event is not contrary to accepted scientific law.

But no scientist is perfect. Scientists sometimes make logical or mathematical blunders, they are not always competent, they are subject to hallucinations, wishful thinking, fatigue. And if a scientist persistently gets results which differ from the general opinion it will be, and is, put down to bias or crankiness or ego-involvement. The worth of a scientist's conclusions is measured against his known 'properties' and any light thrown on these by the vogue psychological theories. Any peculiarities of the observer-calculator are thus granted to be very relevant in any consideration of d_2 .

But although account can be taken in d_2 of the peculiarities or 'accidents' of the observer-calculators, 'essential' properties must remain unacknowledged, *even though they be recognised*. For d_2 cannot be further ameliorated without depriving the notion of 'predictability' of its activity-like characteristics and thus of that component of its content that differentiates it from d_1 , for which 'predicting' will of course signify only a sequence of logical operations given certain boundary conditions. In particular it cannot be stretched to admit one or more classes of events to be unpredictable by any scientist whose perfection leaves him still human: even though many members of a class of events may not in fact be predictable, yet a member of the class must in principle be predictable. D_2 is presently refuted¹ if a scientist² operating under ideal conditions and making the maximum number of relevant and compatible observations, finds himself unable to predict the outcome of an experiment, and this unsucces is repeatable. While the 'in principle' qualification allows the scientist to account for his present failure and promise a future success, by pointing *outwards* at contingently unsatisfactory features in his apparatus or in his observations or in his calculations, it does not allow that he invoke

¹ i.e. relative to the presently accepted theories.

² Or a team of scientists.

DETERMINISM IN CLASSICAL PHYSICS

properties essential to him in his capacity as scientist or man to explain away and excuse perpetual and perhaps inevitable failure. Although therefore at the lower level, the scientist operates with the rich language of scientific theory—which may be as hypothetical as current acceptance decrees, and which, if the theories form a complete lawlike base for the world will be as applicable to the scientist as object, as they are to the events he studies—yet at the level of d_2 , instead of using this same rich language to explain perhaps why not only he, but indeed *any* scientist would be unable to perform certain kinds of predictions, a wholly arbitrary boundary, scientifically-speaking, between accidental peculiarities and essential properties is unjustifiably drawn. Moreover, the first step in a confusing logico-empirico regress in which each step must, on account of the empirical component in a prediction—of predicting as activity—be shorter than the one that precedes it, and in which Gödelian self-references and (predictions)ⁿ leer frightfully from their filigrees, is pointlessly generated.¹

As an attempted scientific explicatum of Determinism, d_2 is radically misconceived. Its assessment and acceptability cannot but rest, as does that of d_1 , on the presently accepted scientific theories. Else it is not scientific.

Yet whereas the evidence for d_1 is simply and non-self-referentially the syntactic form of those theories, that for d_2 is rather the complicated inter-relationship of *all* of those theories with those among them which describe the operations of contingent knowers and calculators. Such involution is foreign to Determinism: Determinism is an assertion about all events, all things, all processes, knowledge and prediction complexes no more, and no less, than others less differentiated. The possibility of the occurrence of certain behaviour patterns, and in particular of successful prediction characteristics, in some of the organisms or systems within and interacting with the World is of no central relevance. That a difference of physical levels sufficient to allow prediction-language to operate may not always be possible, is

¹ Of course, not all of the predictions relevant to establishing d_2 will, even as activities, depend significantly on the essential characteristics of the scientist. But even though we do not recognise the prediction singularities which Popper has claimed to establish for mechanical predictors (see § 4), yet there are at least two analogous types of 'complementarity', in psychoanalysis and in quantum mechanics, in which we would not wish to exclude *a priori* the existence of underlying determinist substrata, even though complete finite predictability may not ever be possible. In any universe in which complementarity (i.e. strong knower-known interactions) was known to occur, d_2 would be distinguishable from d_1 .

a priori no more interesting a fact than, say, the impossibility of perfect heat insulation.

What of the 'necessity' and 'foreordination' of the explicandum? Whereas in d_1 the former is represented as successfully as is possible in these days of non-nomic necessity by the necessity 'implicit' in the deterministic theories, in d_2 even this very much weakened necessity is tainted by 'transmission' through the merely contingent properties of the predictor. 'Foreordination' is indeed explicitly empiricised in d_2 : but attempting to encapsulate the omniscience of divine intuition in a law-determined law-intentional predictor only clouds the clarity of the laws, and in any case does no justice to the intention of Determinism.

4

In this section we attempt to show why Popper's technical thesis is false. This thesis asserts that the proposition:

'For any specified finite prediction task, it is physically possible to construct a predictor¹ capable of carrying out this task'²

is incompatible with classical physics.

By a 'specified finite prediction task' is meant:

the task of predicting, with some chosen and specified degree of precision, some event—i.e. the position and or the velocity of its particles—occurring in a finite mechanical system, sufficiently isolated from outside interference, at a certain chosen future instant of time.²

A summary of the argument runs as follows:

Consider a mechanical system A, not containing a predictor, and a predictor B attempting to predict A. This it can do only if (1) B can calculate the results of its interference with A or if (2) B interferes sufficiently weakly with A.

As to (1), B can assess the way in which it interferes with A either by studying its interfering parts B' and their interaction with A—which means that it has to study the system A + B' instead of A—or on the basis of predictions about itself. The first of these two alternatives does not help for in order to obtain the necessary information about A + B' the whole problem arises again. The second alternative does not help for it will be shown in the next section '8' that a predictor cannot have such information about itself. Thus we must consider (2).³

¹ For a description of a predictor see Popper, op. cit., 123

² Popper, op. cit., 124

³ Popper, op. cit., 128

DETERMINISM IN CLASSICAL PHYSICS

Popper then shows (p. 130 following) that if A *does* consist of a predictor employing the same physical principles as B (or if A and B are of the same 'physical order') (2) is not the case, so that when in '8' self-prediction is shown to be impossible the thesis that, 'if there are classical mechanical predictors . . . then there must exist unpredictable predictors' (p. 133), is apparently established.

We shall first state briefly why we find Popper's argument mistaken, and then describe how a predictor able to perform any specified finite prediction task might in fact be constructed. General criticisms of the standpoint adopted throughout the two papers are that:

(i) The physical construction of the predictors is never made really explicit.

(ii) It is assumed that 'interference' is almost a *tertium quid* coming between predictor and the system to be predicted, requiring apparently separate 'calculation', ((1) above, and hence the necessity of self-prediction), and indeed study as a separate physical object (note the argument following (1) and in particular the assertion that 'the whole problem arises again').

(iii) 'Knowledge', in particular 'self-knowledge' and 'self-information', are given a certain hypostatised subsistence inconsistent with the purely scientific and single-level viewpoint which must be adopted in considering the physical possibility or impossibility of the occurrence of a particular type of sentence token.

This last brings us to the grounds on which we must reject Popper's demonstration, which is that: the sense in which self-prediction is shown to be impossible is not the sense in which self-prediction is necessary in order to allow for the effects of interference. If we distinguish self-prediction *qua* computer, from self-prediction *qua* interaction, then since only the latter is relevant in calculating the effect of interference of the predictor, B, with the system to be predicted, A, Popper's conclusion does not follow, even if we grant that self-prediction *qua* computer is impossible (and '8' is concerned wholly with this). Formally then, if the distinction be allowed Popper's argument is four termed and hence invalid.

But is such a clear distinction between 'computer' and 'interactor' always possible?

It must be understood that in order to disprove Popper's thesis it is only the *possibility* of always being able to construct, for a given specified task, a predictor in which the distinction is embodied and which can therefore perform the task, that we need to establish. It

will of course always be possible to design a 'predictor' which is so confused and inefficient in its working as to be quite incapable of predicting anything, but from the possibility of such bad design nothing follows. However, we shall in our discussion below of the construction of an efficient predictor indicate why the distinction is not only possible but indeed necessary if the property of being a 'predictor' is to be conferred on an otherwise lustreless physical object.

A predictor is a device for making predictions about the future which the designer of the predictor would adjudge to be rational. In order for this to be the case the predictions must be derivable from a theory (together with certain boundary conditions), which is, if the predictions are to be definite and precise, a deterministic theory. If the system A, whose state at some future time t_2 is to be predicted, has f degrees of freedom it is necessary that B, the predictor, should be able to perform two distinct sets of operations:

(i) the measurement of $2f$ independent parameters specifying the state of A (if the equations of motion are second order). The values of these parameters are the initial conditions at time t_1 .

(ii) 'knowing' the f equations governing A's motion it must solve them for the time t_2 , before t_2 has arrived.

For (i) B requires a part B', the interactor, interacting with A through at least f measuring arms (These 'arms' may be mechanical tentacles, thermometers, beams of light, or any other mode of interaction which yields a pointer reading uniquely determined by the conditions obtaining at the place of interaction). If the effect of B' on A is small relative to the degree of precision required in the predictor it can be left in continuous contact with A without ill effect. Most classical measurements are of this nature and, giving rise to no prediction difficulties, correspond to what Popper has called 'electromagnetic miracles'. For (ii) B requires a part B'', the computer.

We shall consider only the case in which the interaction between B' and A is strong, and our predictor will operate in the following manner.

(a) At time t_1 B' is placed in contact with A for a time Δt_a seconds after which time it is withdrawn from contact in a determinate manner.

(b) Δt_b seconds after withdrawal from A, B' communicates to B'', the computer, $2f$ pointer readings virtually instantaneously and is then disconnected. In this communication no *physically* important interaction occurs; any attenuation, etc., in the pointer readings of B' during the communication can be allowed for in B''s interpretation of these readings, provided always determinateness, i.e. uniqueness, obtains.

DETERMINISM IN CLASSICAL PHYSICS

(c) B'' is now successively programmed to have a logical structure isomorphic with a method of solution of: the equations governing the motion of B' alone (solved retrodictively for a period of time Δt_b , from $t_1 + \Delta t_a + \Delta t_b$ to $t_1 + \Delta t_a$); the equations relating the $2f$ independent parameters of B' to the $2f$ independent parameters of A (those will be determined by the structure of the composite system $A + B'$); and finally, the equations of motion of A alone (solved predictively for the period $t_1 + \Delta t_a$ to t_2). Suppose that the whole calculation takes Δt_c seconds. The resultant prediction is produced, for example, in the form of a punched tape.

We must consider several points in this account in greater detail.

All measurement involves a physical interaction and a consequent disturbance of the measured system. If A is a very delicate system it may be the case that any B' we can design interacts so strongly with A during the period Δt_a of contact as to distort seriously or even to transform the future of A . However this has no bearing whatsoever on whether we can predict that future. So long as the interesting singularities of the Oedipus effect¹ are avoided, i.e. so long as the prediction is a logical deduction rather than a physical interactant and causal determinant, the fact that knowledge or information can only be obtained at the price of interference is a fact which is no harder than most facts.

In order that a prediction should be a scientific prediction (i.e. a prediction with rational backing) and not just a prophecy, it is necessary that it be logically deducible from certain other propositions which make some sort of claim to some sort of truth. The form of a deterministic theory requires that certain initial conditions be given once and for all, and then that certain mathematical equations be solved in a logical manner. All well-designed predictors applying determinist theories must accordingly be constructed so that B'' the computer is (during the computation) 'logically isolated' from A and B' once the initial communication of information has taken place. By 'logical isolation' is meant: the physical events in B'' must, in so far as they possess logical or linguistic significance, be physically independent of any events occurring in A and B' . For if B'' is to be a physical embodiment of certain logical and mathematical operations the only physical events occurring in B'' and affecting the final prediction must be events

¹ See Popper, *op. cit.*, 188. And this effect *must* be avoided (or be negligible) in any scientific prediction which is to be acted upon by rational men (who will not be unconscious of this fact too).

possessing a propositional interpretation. Hence the distinction between predictor qua computer and predictor qua interactor is not only possible in fact, as we shall see almost at once, but is logically necessary.

The simplest way of achieving 'logical isolation' is to make B' and B'' physically quite distinct as suggested above. In certain cases (i.e. for certain A's), this physical isolation may not be possible, but even then the interference must be present only 'as a perturbation', and problems of self-reference and feedback circumvented if the logical character of the prediction is to be preserved.

But even granting that such 'logical isolation' is logically necessary is it always—for all specified prediction tasks—physically possible?

Since it would appear to be necessary for all predictors to be amplifiers¹, and since most modern computers are digital, the most convenient way of making contact between B' and B'' would be by electronic means. This would have two advantages. Firstly, if all B''s readings are expressed as fluctuating potentials on the grid of a valve the reaction of B'' on B' can be made very small (so that the 'attenuation, etc.' mentioned in (b) above would be unimportant because electronically an 'electromagnetic miracle', or the one-way passage of information, is always possible; note, however, that a finite reaction would not affect the argument). Secondly 'logical isolation' can be very simply achieved by 'biasing off' the input valves after the initial communication of information. Although electronic devices would appear to be applicable to all classical prediction tasks (so that the question above is answered affirmatively) the possible use of electronics is not essential to this argument. Temporary mechanical contact between B'' and a B' disconnected from A is an alternative (e.g. for analogue computers).

This concludes the discussion of the construction of an efficient predictor. Basically it has centred round the fact that any definite or uniquely determinate interaction with a physical system is potentially interpretable as a measurement of that system, and that if the observed system possesses f degrees of freedom any $2f$ independent pointer indications will, no matter how processed, enable its future to be predicted provided that they somehow filter through to a logically isolated computer in a determinate manner.

Let us now suppose that the system A is itself a predictor. Can the predictor B predict A's future state?

¹ Popper, *op. cit.*, 129

DETERMINISM IN CLASSICAL PHYSICS

If A is itself predicting a system X the problem is simple. If B happens to be a more efficient predictor than A it can, by looking at A's amplified input from X (and this need not disturb A or X) and operating a programme based on A's logical structure when it is reacting to X, predict A's future behaviour and in particular A's prediction of X. But if B is slower in its computations than is A it will not be able to predict A's future prediction for the very good physical reason that

$$\Delta t_a + \Delta t_b + \Delta t_c > t_2 - t_1 \text{ (See (a), (b), (c) above)}$$

More interesting is the possibility of prediction when A is studying B and B is studying A, and A and B are of the same physical order (i.e. the interaction is strong). This is the problem with which Popper is initially concerned.

If A is to qualify as one of Popper's 'finite prediction tasks' there are two and only two possibilities.

(a) B is required to predict only A', the interactor. Both the construction of A' and the rules by which the feelers of A' are instructed (by A'') to search out its surrounds must be programmed into B, which can then perform its task. For this case it would be possible both for B to predict A' and A'' to predict B'. Since the interaction is strong A' and B' will exert a strong influence on one another's futures.

(b) B is required to predict the future not only of A' but of A'' too. In order to do this all of A's programmes, together with instructions about their use must be known by B, and further, B'' must be a faster computer than A''.¹ It is logically impossible to have B totally predicting A, and A totally predicting B without a solution of the problems of self-reference,² but this has no bearing on B's ability to predict any *specified* physical system A, and so does not concern us.

¹ How such knowledge is obtained by B is really irrelevant to our problem in which we are concerned only with *specified* A's. The knowledge could be programmed into B by the makers, or obtained by B studying A prior to the actual prediction task. During this prior examination it would be necessary that: (i) A be incapacitated in its function (if it has one) of self-programming computer (B, having cosmic aspirators, might be programmed to remove, on occasion, the power supply to any A); or (ii) B was able to arrive at a satisfactory hypothesis explaining A and its self-programming potential *in toto*.

² The problem of self-reference as it occurs in Popper (*loc. cit.*) is: given that S is the case is it possible to construct a language L such that 'S' describes S + "S" where "S" is the sentence token for 'S'. Given a *finite* number of possible states (S₁S₂ . . . S_n) there seems to be no reason why a suitable L should not be so engineered.

We therefore find that the only obvious limitations imposed by virtue of the physical nature of predictive activities on the prediction of A by B are due to:

- (i) the finite operating speed of the computer and of measurement (requiring $t_2 - t_1 > \Delta t_a + \Delta t_b + \Delta t_c$),
- (ii) A not being sufficiently large to interact with some B',
- (iii) the physical impossibility of B learning of the construction of A.

Neither (ii) nor (iii) are the case if A is itself a predictor and is a finite prediction task, and (i) is a physical requirement devoid of any logical significance.

We conclude that predictors are in principle predictable, and so that there is therefore no logical reason why in a determinist universe any system should not be predictable.

Emmanuel College
Cambridge

A CAUSAL CALCULUS (I) *

I. J. GOOD

I Introduction

THIS paper contains a suggested quantitative explication of probabilistic causality in terms of physical probability.¹ The main result is to show that, starting from very reasonable desiderata, there is a unique meaning, up to a continuous increasing transformation, that can be attached to 'the tendency of one event to cause another one'. A reasonable explicatum will also be suggested for the degree to which one event caused another one. It may be possible to find other reasonable explicata for tendency to cause, but, if so, the assumptions made here will have to be changed.

I believe that the first clear-cut application in science will be to the foundations of statistics, such as to an improved understanding of the function of randomisation, but I am content for the present to regard the work as contributing to the philosophy of science, and especially to what may be called the 'mathematics of philosophy'. Light may also be shed on problems of allocating blame and credit. I hope to consider applications to statistics on another occasion.²

In a previous note³ I have tried to give an interpretation of 'an event F caused another event E' without making reference to time. It was presumably clear from the last three paragraphs, which were added in

* Received 19. I. 1960

¹ Compare : (i) Hans Reichenbach, 'Die Kausalstruktur der Welt und der Unterschied von Vergangenheit und Zukunft', *Ber. d. Bayer. Akad. d. Wissensch., math.-nat. Abt.*, 1925, pp. 133-175; or Chapter 3 of his book *Modern Philosophy of Science*, London and New York, 1959; (ii) Norbert Wiener, 'The theory of prediction', in *Modern Mathematics for the Engineer* (ed. by E. F. Beckenbach), New York, 1956, pp. 165-190. The present work goes further than Reichenbach's in offering a definite explicatum. It bears little resemblance to Wiener's work, which is mathematically much more advanced.

² The present paper owes much to correspondence and discussion with Mr E. M. L. Beale, Professor Bruno de Finetti, Professor K. R. Popper, Professor L. J. Savage, Mr Christopher Scott, and especially to Dr Oliver Penrose. The Referees and Editor have also been helpful.

³ This *Journal*, 1959, 9, 307-310

proof,¹ that I was not satisfied with my attempt.² I shall describe the note as the 'previous paper' but it will not be necessary for the reader to refer back.

The present paper is more ambitious in that it is quantitative, but less so in that it assumes, at least at first, that F is earlier than E. (It may be possible to interpret the explicatum more generally.) As in the previous paper I shall take for granted the notion of physical (= material) probability. In order to avoid misunderstanding I must mention my opinion that in so far as physical probability can be measured it can be done only in terms of subjective probability, but this opinion will not affect the arguments below. Likewise the notion of an 'event' will be taken for granted. In some earlier drafts I included material dealing with the meanings of 'event', 'probability', and 'definition', and with the modifications of the analysis required in order to cope with the possibility that the future may affect the past. I have omitted this material here for the sake of brevity.

2 Notation and General Outline

Propositions and events will be understood in a very general sense, and will be denoted by the symbols E, F, G, H, and U. These will be combined by means of the logical connectives '.' meaning 'and', '-' meaning 'not', and 'v' meaning 'or'. A vertical stroke, '|', will mean 'given', as in the expression $P(E | H)$, the probability of E given H. Similarly $O(E | H)$ will mean that the odds of E given H, i.e. $P(E | H)/P(\bar{E} | H)$. Sometimes some or all of what is 'given' is omitted from the notation for the sake of brevity. A colon will be used to mean 'provided by' or 'by' or 'from', as in

$$I(E : F | G) = \log(P(E | F \cdot G)/P(E | G)),$$

which can be read from left to right as the amount of information concerning E provided by F given G. Another example of the colon notation is

$$\begin{aligned} W(H : E | G) &= \log(P(E | H \cdot G)/P(E | \bar{H} \cdot G)) \\ &= \log(O(H | E \cdot G)/O(H | G)) \\ &= I(H : E | G) - I(\bar{H} : E | G), \end{aligned}$$

¹ The words 'added in proof' were omitted in error, and the effect was peculiar.

² I find that Reichenbach made a similar error when defining the expression 'causal relevance'. See Appendix I.

A CAUSAL CALCULUS

the 'weight of evidence concerning H provided by E given G'.¹

The general plan of the paper is to suggest explicata for:

(i) $Q(E : F)$, or Q for short, the 'causal support for E provided by F, or the tendency of F to cause E'. The explicatum that the argument forces upon us is the weight of evidence against F if E does not happen, $W(\bar{F} : \bar{E})$, or more explicitly, $W(\bar{F} : \bar{E} \mid U \cdot H)$, where U and H are defined below. In order to formulate enough desiderata it is necessary to introduce some auxiliary notions.

(ii) The strength of a causal chain joining F to E.

(iii) The strength of a causal net joining F to E. (Causal chains and nets will be defined in Sections 8 and 11.)

(iv) $\chi(E : F)$, or χ for short, the contribution to the causation of E provided by F, i.e. the degree to which F caused E. This will be defined as the strength of a causal net joining F to E, when the details of the net are completely filled in, so that there are no relevant events omitted. (I avoid the use of the letter C for either Q or χ , because it has been used to mean corroboration.²) An example is given in Appendix II to show that Q and χ cannot be identified.

It would not be appropriate to define χ as the limit of strengths of more and more detailed nets; for, if space and time are continuous, the limiting operation could be done in a biased manner so as to get entirely the wrong result; like a lawyer making a case by special selection of the evidence. We must have the whole truth in order to define χ in principle. (Compare the first Appendix.) If, however, the events fill the relevant parts of space and time, and we let the events become smaller and smaller, then the limit should be unique.

In practical uses of the notion of causality, judgments of approximate irrelevance are always made in order to reduce the complication of the causal net.

It is possible to draw an analogy between a causal net and an electrical resistance network, with a resistance in each link. In this analogy it is necessary to imagine a rectifier placed in each link in order to prevent a flow of causal influence backward in time. It is then tempting to define the degree of causality between the input and output of

¹ See, for example, 'Weight of Evidence, Corroboration, Explanatory Power, Information, and the Utility of Experiments', *J. Roy. Stat. Soc., ser. B*, 1960, 22, 319-331, where there are further references.

² Karl R. Popper, *The Logic of Scientific Discovery*, London, 1959, pp. 379-418, mostly published in this *Journal*, 1954, 1957, and 1958. See also Note 1 above.

the causal net as the effective resistance of the corresponding causal network. This analogy suggests that the causal resistances should be defined in such a manner that they are additive for chains in 'series', and such that their reciprocals are additive for chains in 'parallel'. It turns out that the analogy cannot be pressed as far as this, but it is one of the themes of the paper.

The main part of the paper consists of a list of assumptions, and deductions from them, leading up to the above explicatum for Q . Afterwards a general explicatum will be suggested for χ , but this will not be deduced in the same formal manner. It is fairly convincingly unique for causal nets of the 'series-parallel' type, and has a certain cogency in the general case.

3 *Small Events*

Until near the end of the paper all events will be assumed to occupy small volumes of space (more precisely : have small diameters) and occupy small epochs of time. For the most part 'space' could be interpreted in a more general sense than as ordinary three-dimensional space ; for example, it could be phase space or Hilbert space. On the other hand time will be assumed to be well-ordered and one-dimensional. There must be some sort of metric in both space and time, in order that smallness and contiguity should have a meaning. If the metrics of space and time have been mixed up, as in the theory of relativity, then they will be assumed to be unmixed by the use of a fixed frame of reference. (Dr O. Penrose has pointed out that the present work is consistent with the theory of relativity provided that causal influence does not travel faster than light.)

4 *Causal Support, or Tendency to Cause*

Let H denote all true laws of nature, whether known or unknown, and let U denote the 'essential physical circumstances' just before F started. When we talk about 'essential physical circumstances' we imply that the exact state has a probability distribution. An equivalent description is to say that a system is one of an 'ensemble'. (I must admit that there is more than meets the eye in this description, since in quantum mechanics the word 'state' can be given at least eight interpretations, seven of which may be relevant here. See Appendix 3.)

A CAUSAL CALCULUS

In order to arrive at explicata for Q and χ I have found it necessary to discuss them in an interconnected manner ; i.e. there do not appear to be enough desiderata for Q , considered by itself, to circumscribe its possible explicata to a satisfactory extent.

In the present section the ground will be cleared by discussing what Q and χ depend upon. It is convenient to think of this dependence in terms of notation, which seems to bring out the main points better than a purely verbal discussion. For example, the symbols Q and χ are abbreviations for $Q(E : F)$, and $\chi(E : F)$, and these expressions are themselves abbreviations for $Q(E : F | U . H)$ and $\chi(E : F | U . H)$. To take U and H for granted, and omit them from the notation, is parallel to linguistic usage. If we say that it is bad for eggs to throw them in the air, we take it for granted that there is a law of gravitation, and that there is a large gravitational body nearby.

Events later than F and earlier than E may be relevant to χ but not to Q . Accordingly I shall assume that $Q(E : F)$ depends only on $P(E | F)$, $P(E | \bar{F})$, $P(E)$, and $P(F)$. It is natural to define $Q(E : F | G)$ as the same function of these four probabilities, but made conditional on G .

Even $Q(E : F | U . H)$ is an incomplete notation. If the subjective element is to be removed from the expression ' F caused E ', then it must be expanded to ' F , as against \bar{F}_D , caused E rather than E' ', where the suffix, D , to \bar{F} (the negation of F), represents a complete specification of the relative probabilities of the mutually exclusive events whose disjunction is \bar{F} . (D represents a probability distribution.) We could use a notation like

$$Q(E/E' : F/\bar{F}_D | U . H)$$

or

$$Q(E : F/F_D | U . H . (E \vee E')),$$

the degree of causation of E rather than E' by F rather than \bar{F}_D .

The failure to recognise all the variables on which tendency to cause is based was for me one of the stumbling blocks in capturing the notion of probabilistic causality, if indeed I have fully succeeded even now.

It should be held in mind that $E \vee E'$ is regarded as taken for granted in the four probabilities on which Q is assumed to depend, when we are concerned with the causation of E rather than E' . When we take $E \vee E'$ for granted we may write \bar{E} instead of E' .

5 *Assumptions and Deductions Leading to the Explicatum for Q*

In order to make my assumptions clear I shall list them in the form of axioms, A_1, A_2, \dots ; and the deductions from them will be called theorems T_1, T_2, \dots , for ease of reference. On a first quick reading the justifications and proofs should *quite definitely* be skipped, but I have not postponed them to a later section. (I did so in an earlier draft, but the cross-referencing made the paper more difficult to read.) The justifications of the most easily acceptable axioms, and the proofs of the easily proved theorems will be omitted.

Let $P(F) = x$, $P(E | F) = p$, $P(E | \bar{F}) = q$, $P(E) = r$, so that

$$r = xp + (1 - x)q, x = (r - q)/(p - q).$$

Unless $p = q$ (in which case $r = p = q$), x is a function of p, q , and r . Therefore by an assumption of the previous section we have :

A_1 . $Q(E : F | G)$ is a function of p, q, r , unless perhaps $p = q = r$. We call this function $Q(p, q, r)$ so that the symbol Q has two slightly different meanings. The symbol G will usually be taken for granted and omitted.

A_2 . Q is a real number or ∞ or $-\infty$; but it may be indeterminate for special values of p, q , and r , such as when two of them are equal or one of them is 0 or 1. (It seems sensible to look for a scalar explicatum rather than a 'vector'.)

The next two axioms deal with monotonicity and continuity.

A_3 . (i) Q increases with p if q and r are held constant.

(ii) Q decreases when q increases if p and r are held constant.

A_4 . Q is continuous except when it becomes infinite or indeterminate, if there are such points.

A_5 . If $P(F) \neq 1$, meaning, as usual, $P(F | U \cdot H) \neq 1$, then Q has the same sign as $p - r$, and therefore also the same sign as $p - q$, and as $r - q$; and if these expressions vanish we may say that F has no tendency to cause E , and we put $Q = 0$. (This axiom removes a gloss from A_1 .)

A_6 . Any causal net joining F to E , as defined below in Section 11, has a causal strength, S , and a causal resistance, R . These are positive numbers, except that if $p = q = r$, or if p or q is 0 or 1, we may get zero or infinite resistance or strength. (An important part of the definition of a causal net is that it consists only of events that actually occurred or will have occurred.)

A_7 . There is a functional relationship between R and S , $S = f(R)$, $R = g(S)$, where f and g are absolute functions inverse to one another.

A CAUSAL CALCULUS

A8. *f and g are continuous decreasing functions.*

A9. $\chi(E : F)$ is the strength of the complete causal net joining F to E. More precisely, it is the limit, as the sizes of the events tend uniformly to zero, of the strengths of nets ; where each net of the sequence joins F to E, consists of a finite number of events, and omits no events temporarily between F and E. It is not claimed that this axiom is formulated with complete rigour, but it is used in only a weak form for the explication of Q (in the proof of T14). It is introduced at this early stage in order to supplement the outline in Section 2. (If we assume that the degree to which F caused E has an objective meaning, with a precise numerical value, we are committed to the idea that there is a complete world, uninfluenced from outside. Outside influence could be allowed for by assuming that the numerical values are not absolutely precise.)

A10. *The strength of the causal net consisting of F and E alone is equal to $Q(E : F)$ when this is positive, and is otherwise zero. (We shall clearly get nowhere unless we assume some relationship between Q and S, and A10 is the simplest reasonable relationship that could be assumed.)*

The strength and resistance of a net, n , joining F to E, are denoted by $S(E : F | n)$ and $R(E : F | n)$.

A11. *Let n be a 'chain', $F = F_0 \rightarrow F_1 \rightarrow \dots \rightarrow F_n = E$. Then $S(E : F | n)$ is some function, $\phi(s_0, s_1, \dots, s_{n-1})$, of the strengths, s_0, s_1, \dots, s_{n-1} , of the links. Here*

$$s_i = S(F_{i+1} : F_i | U_i \cdot H),$$

where U_i represents the essential physical circumstances just before F_i began. (Causal chains are formally defined in Section 8.)

T1. $\phi(s) = s$. (Proof from A10 and A11.)

A12. ϕ is a symmetrical function, i.e. the arguments of the function can be permuted without changing its value.

A13. ϕ is a non-decreasing function of each of its arguments. (A chain cannot be weakened by strengthening any of its links without changing the strength of any of the others.)

T2. ϕ vanishes if the chain is cut, i.e. if any of the links is of zero strength. We may alternatively say in this case that there was no causal chain.

A14. *If two consecutive links are replaced by a single link of equivalent strength, then the strength of the chain is unchanged. Formally,*

$$\begin{aligned} &\phi(s_0, s_1, \dots, s_{n-1}) \\ &= \phi(s_0, s_1, \dots, s_{i-1}, \phi(s_i, s_{i+1}), s_{i+2}, \dots, s_{n-1}). \end{aligned}$$

A15. *A chain is not weakened by 'omitting' one of its links, i.e.*

$$\begin{aligned} &\phi(s_0, s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_{n-1}) \\ &\geq \phi(s_0, s_1, \dots, s_{n-1}). \end{aligned}$$

T3. *A chain is no stronger than its weakest link.* (From A10 and A15.)

Definition. Let the maximum possible causal strength be σ . This is either a positive finite number or $+\infty$.

A16. $S(E : F | n) \leq \sigma$, for any net, n . (This axiom is a mere restatement of the definition.)

T4. $Q(p, q, r) \leq \sigma$. (From A10 and A16.)

A17. *If any of the links of a chain is of strength σ , then it can be 'omitted' in the sense of A15, without strengthening the chain.*

$$\begin{aligned} T5. \quad & \phi(s_0, s_1, \dots, s_{i-1}, \sigma, s_{i+1}, \dots, s_{n-1}) \\ &= \phi(s_0, s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_{n-1}). \end{aligned}$$

(From A15 and A17.)

T6. *If every link of a chain is 'cast-iron', then the chain is cast-iron, i.e. $\phi(\sigma, \sigma, \dots, \sigma) = \sigma$.* (From T1 and T5.)

A18. *A chain of n links all of the same fixed strength, τ , where $\tau < \sigma$, is as weak as you like if n is large enough. Formally $\phi(\tau, \tau, \dots, \tau) \rightarrow 0$ as the number of arguments tends to infinity.*

A19. ϕ is a continuous function of all its arguments when they are all less than σ ; and, if $s_i \rightarrow \sigma$, then the value of the function tends to the value it would have with $s_i = \sigma$. The reason for the clumsy expression of this axiom is that σ may be $+\infty$.

T7. *If a chain has n links, all of the same strength, τ , where $\tau < \sigma$, then the chain is as strong as you like if n is fixed and τ is close enough to σ . Formally, if n is fixed, then*

$$\phi(\tau, \tau, \dots, \tau) \rightarrow \sigma \text{ when } \tau \rightarrow \sigma.$$

(From T6 and A19.)

T8. *There is a function, g , such that, identically,*

$$\phi(s_0, s_1, \dots, s_{n-1}) = g^{-1}(g(s_0) + g(s_1) + \dots + g(s_{n-1})).$$

The function, g , is defined for all non-negative arguments not exceeding σ , and is itself non-negative, continuous and strictly decreasing, and $g(0) = \infty$, $g(\sigma) = 0$. We may define g as $+\infty$ when its argument is negative.

Proof. Consider the function $\phi(s, t)$ of just two variables. By A19, A13, A12, and A14, this function may be said to be continuous, monotonic, commutative, and associative. It follows that it is of the form $g^{-1}(g(s) + g(t))$, where g is a strictly monotonic continuous function. The use of the symbol g is justified since A7 and A8 can be satisfied with this function.

A CAUSAL CALCULUS

The mathematical theorem just invoked was apparently first published by Abel.¹ It was rediscovered several times, such as by Aczél.² What it amounts to is that ϕ can be calculated on a suitably calibrated slide-rule.

The results $g(0) = 0$, $g(\sigma) = \infty$, follow from A18 and A19.

Q. E. D.

We may satisfy A6, A7, and A8, which are the only axioms that involve R, by identifying $g(S)$ with R. This identification is no restriction on the explication of Q. As a consequence of this identification we have the following theorem.

T9. *The resistance of a chain is equal to the sum of the resistances of its links.*

A20. *Consider the causal net shown in the diagram below, in which $P(F) = x$, $P(G_j | F) = p_j$, $P(G_j | \bar{F}) = q_j$, $P(G_j) = r_j = xp_j + (1-x)q_j$; $P(E | G_1 \vee G_2 \vee G_3) = 1$, $P(E | \bar{G}_1 \cdot \bar{G}_2 \cdot \bar{G}_3) = 0$, $P(E) = r$, $j = 1, 2, 3$, and where G_1, G_2, G_3 are independent given F and also given \bar{F} . Then the strength of the net is a function of the strengths of the three separate chains, and this function is continuous, monotonic increasing in each variable, commutative (cf. A12), and associative (cf. A14).*

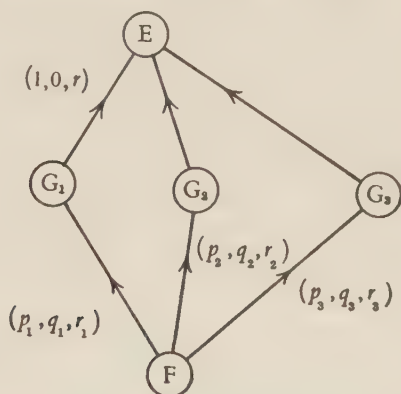


FIG. 1

T10. *The strength of the net of A20, generalised to m chains in parallel, is of the form*

$$h^{-1}(h(s_1) + h(s_2) + \dots + h(s_m)),$$

¹ Neils Henrick Abel, *Oeuvres Complètes*, tome 1, 1881. The paper was Abel's first publication.

² J. Aczél, *Bull. Soc. Math. France*, 1948, **76**, 59-64. See also Aczél, *Acta Phys. Acad. Sci. Hungar.*, 1955, **4**, 351-362; L. Jánosy, *loc. cit.* pp. 333-349. Or see *Math. Rev.*, **10**, 685, and **16**,¹¹²⁷⁻¹¹²⁸.

where s_1, s_2, \dots, s_m are the strengths of the individual chains. The function h is defined for all non-negative arguments not exceeding σ , and is itself non-negative, continuous and strictly increasing, and $h(0) = 0$, $h(\sigma) = \infty$. (The theorem of the generalised slide-rule implies this theorem, just as it implied T8 above.)

We are now at liberty to call $h(S)$ the strength of a causal net, in place of S , provided we are content to determine the explicatum of S and Q only up to a continuous increasing transformation. It might be thought for a moment that this change of notation would invalidate T9. But since T8 is now true with $g(x)$ replaced by $g(h^{-1}(x))$, we can simply rename this function ' $g(x)$ ' in order to validate T9. With these conventions we have:

T11. *The strength of the net of A20, generalised to m chains in parallel is the sum of the strengths of the individual chains.* When applying this theorem the independence condition mentioned in A20 should not be overlooked.

It appears that the analogy with electric networks is not bad, although the function $f(x)$ turns out later not to be $1/x$, but another self-inverse function.

A21. *In the net of A20, with G_3 omitted, i.e. with only two chains in parallel, we may regard $G = G_1 \cdot G_2$ as a single event, without changing the causal strength of the net.* Note that it would be unreasonable to assume this coalescence property for dependent events, for if we did so we could collapse any net into a single event.

It may be objected that $G_1 \cdot G_2$ is not necessarily a small event. But the strength of a causal net should depend only on its topology, with time-order preserved, and on the various probabilities and conditional probabilities. Hence $G_1 \cdot G_2$ could be a small event in an equivalent network.

Although I think A21 is eminently reasonable, especially in view of later developments, as in Section 9, I believe it to be the weakest part of my argument, and I conjecture that the replacement of this axiom by other assumptions would be the most fruitful method of finding other explicata of tendency to cause, if they exist.

T12. *Identically, if $p_1 \geq q_1$, $p_2 \geq q_2$, $0 < x < 1$, then*

$$S(p_1 + p_2 - p_1 p_2, q_1 + q_2 - q_1 q_2, x(p_1 + p_2 - p_1 p_2) + (1 - x)(q_1 + q_2 - q_1 q_2))$$

$$= S(p_1, q_1, x p_1 + (1 - x) q_1) + S(p_2, q_2, x p_2 + (1 - x) q_2).$$

This follows at once from T11 by making the identification mentioned in A21.

A22. $Q(p, q, r)$ is an analytic function when $0 < p < 1$, $0 < q < 1$, $0 < r < 1$, $p \neq q$.

The only purpose of this axiom is to enable us to extend a formula proved for a large set of values of (p, q, r) to all values except those for which Q may be infinite or indeterminate. I think only an extreme purist would object to A22. It could be avoided by assuming instead that Q is anti-symmetric in the sense

$$Q(p, q, xp + (1 - x)q) = -Q(q, p, xq + (1 - x)p).$$

T13.

$$Q(p, q, r) = u(x) \log \frac{1 - q}{1 - p} = u \left\{ \frac{r - q}{p - q} \right\} \log \frac{1 - q}{1 - p},$$

where $u(x)$ is a non-negative analytic function of x .

This theorem, and the next one, will be superseded by T15.

Proof. By A10 and A22, we may replace S by Q in T12, and drop the inequalities $p_1 \geq q_1$, $p_2 \geq q_2$. Let

$$\psi(\xi, \eta, x) = Q(1 - e^\xi, 1 - e^\eta, x(1 - e^\xi) + (1 - x)(1 - e^\eta)),$$

$$p_1 = \exp \xi_1, q_1 = \exp \eta_1, \text{ etc.}$$

Then

$$\psi(\xi_1 + \xi_2, \eta_1 + \eta_2, x) = \psi(\xi_1, \eta_1, x) + \psi(\xi_2, \eta_2, x).$$

On putting $\eta_1 = \eta_2 = 0$, and provisionally regarding x as a constant, we get a well known functional equation whose only continuous solution is easily seen to be of the form

$$\psi(\xi, 0, x) = \xi \cdot u(x),$$

where $u(x)$ is a function of x only. (The only other solutions are in fact non-measurable.¹) Likewise $\psi(0, \eta, x) = \eta \cdot w(x)$, where $w(x)$ is a function of x only. Therefore

$$\begin{aligned} \psi(\xi, \eta, x) &= \psi(\xi + 0, 0 + \eta, x) \\ &= \psi(\xi, 0, x) + \psi(0, \eta, x) = \xi u(x) + \eta w(x). \end{aligned}$$

Therefore

$$Q(p, q, xp + (1 - x)q) = u(x) \cdot \log(1 - p) + w(x) \cdot \log(1 - q).$$

T13 now follows from A5 combined with the equation

$$r = xp + (1 - x)q.$$

Q. E. D.

A23. Consider a radioactive particle in a certain state, which I shall call the 'white' state. In any time interval, t , it has probability $e^{-\alpha t}$ of remaining in the white state throughout the interval if it starts the interval in that state.

¹ G. Hamel, 'Eine Basis aller Zahlen und die unstetigen Lösungen der Funktionalgleichungen $f(x + y) = f(x) + f(y)$ ', *Math. Annalen*, 1905, **60**, 459-462. Or see G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, Cambridge 1934, p. 96.

If it does not remain in the white state, then it proceeds to another state called here the 'black' state, from which there is no return. Now let F be the event that the particle is in the white state at the start of an interval of duration T and let E be the event that it is in the white state at the end of this interval. Then we assume that, if F and E both occurred, $\chi(E : F)$ does not depend on the unit in terms of which time is measured.

A24. If $F \cdot E$ implies G , and $F \rightarrow G \rightarrow E$ is a chain, then this chain is of the same strength as $F \rightarrow E$.

T14. $R(p, o, r) = \nu(r/p) - k \cdot \log p$,
where $\nu(x)$ is a non-negative analytic function of x , and k is a positive constant.

Proof. Consider the radioactive particle described in A23. Let $P(F) = x$. The degree to which F caused E is the limit of the strengths of finite chains obtained by breaking up the time interval (o, T) into a 'Riemann dissection' (see A9). Since g is a continuous function (A8) the resistances of these finite chains must also tend to a limit, which we may call the causal resistance from F to E . This must be some function of x , α , and T , say $R^*(x, \alpha, T)$. By A23 we see that for any positive constant, k , the resistance must be equal to $R^*(x, k\alpha, T/k)$. Since this is independent of k it must be of the form of $R^*(x, \alpha T)$.

Now, by a continuity argument, we may generalise T9 to continuous chains, and hence deduce that, for any positive T and U we have

$$R^*(x, \alpha T) + R^*(I, \alpha U) = R^*(x, \alpha T + \alpha U).$$

By giving x the value 1 and subtracting from the equation with arbitrary x , we see that $R^*(x, \alpha T)$ is of the form

$$R^*(x, \alpha T) = \nu(x) + R^*(\alpha T),$$

where, identically,

$$R^*(\alpha T_1 + \alpha T_2) = R^*(\alpha T_1) + R^*(\alpha T_2),$$

so that $R^*(\alpha T)$ is of the form

$$R^*(\alpha T) = k_1 \alpha T.$$

Now, by repeated use of A24, we see that

$$R(p, o, xp) = R^*(x, \alpha T),$$

where $p = e^{-\alpha T}$. Thus

$$R(p, o, r) = \nu(r/p) - k \cdot \log p.$$

Q. E. D.

$$T15. \quad Q(p, q, r) = \log(1 - q) - \log(1 - p),$$

$$R(p, o, r) = -\log p,$$

where the base of the logarithms may be taken as e . $Q(p, q, r)$ is mathematically independent of r , and may be abbreviated to $Q(p, q)$. It can be written in other ways :

A CAUSAL CALCULUS

$$\begin{aligned} Q(E : F | G) &= \log \frac{P(\bar{E} | \bar{F} \cdot G)}{P(\bar{E} | F \cdot G)} = \log \frac{O(\bar{F} | \bar{E} \cdot G)}{O(\bar{F} | G)} \\ &= W(\bar{F} : \bar{E} | G) = -W(F : E | G), \end{aligned}$$

the weight of evidence against F if E does not happen. More precisely, Q is uniquely determined only up to a continuous analytic increasing transformation. Among all the explicata there is just one apart from a scale factor (choice of unit), for which theorems T9 and T11 are true. We lose no real generality, and we gain simplicity, by choosing this explicatum.

Proof. By T13, T14, and A7, we have the identity

$$f(v(r/p) - \log p) = -u(r/p) \cdot \log(1 - p).$$

Let $v(x) = \gamma$, $-\log p = z$, $\log f(\gamma + z) = \rho(\gamma + z)$. Then $\rho(\gamma + z)$ is of the form

$$\rho(\gamma + z) = \rho_1(\gamma) + \rho_2(z).$$

If $v(x)$ is not a constant, we can differentiate and deduce that $\rho(\gamma)$ is a linear function of γ , from which we can soon derive that $\log(1 - p)$ is a power of p . Since this is false it follows that $v(x)$ is a constant, and hence also that $u(x)$ is a constant.

The theorem now follows from the remark that the choice of the base of the logarithms is equivalent merely to the choice of units of measurement of strength and resistance. We may call the units 'natural', 'binary', or 'decimal', according as the base is e , 2, or 10. In this paper I shall use natural units. Possible names would be 'natural causats' and 'natural tasuacs'.

Note that the explicatum for Q was by no means obvious in advance, nor was it obvious that all the desiderata could simultaneously be satisfied.

It is interesting to note that, if, contrary to most of the discussion, we assume E to be earlier than F , and if the universe has the 'Markov' property, defined below, then the tendency of F to cause E is zero. This result may very well have been taken as a desideratum, but was in fact noticed only after the explicatum was obtained. By the Markov property is meant here that, for prediction, a complete knowledge of the immediate past makes the remote past irrelevant.

T16. *The relationship between R and S is symmetrical, namely*

$$\begin{aligned} R &\geq 0, S \geq 0, \\ e^{-R} + e^{-S} &= 1, \end{aligned}$$

or equivalently,

$$R = -\log(1 - e^{-S}), S = -\log(1 - e^{-R}).$$

Further,

$$R(p, q, r) = \log (1 - q) - \log (p - q).$$

This is an immediate corollary of A7 and T15.

Thus the function f is its own inverse, g . It is tempting to permit negative and imaginary values because some of the formalism is faintly reminiscent of Feynmann's formulation of quantum mechanics, but I shall not pursue this matter here.

T17. If a chain consists of n links whose p 's and q 's are (p_i, q_i) , where $p_i \geq q_i$, then its causal strength is

$$- \log \left\{ 1 - \prod_i \frac{p_i - q_i}{1 - q_i} \right\}.$$

This follows from T16 and T9.

Before reading the proofs in the present section the reader will probably prefer to read the next two sections, in which some examples are given.

Appendix. Holmes, Moriarty, and Watson (see section 2)

Sherlock Holmes is at the foot of a cliff. At the top of the cliff, directly overhead, are Dr Watson, Professor Moriarty, and a loose boulder. Watson, knowing Moriarty's intentions, realises that the best chance of saving Holmes's life is to push the boulder over the edge of the cliff, doing his best to give it enough horizontal momentum to miss Holmes. If he does not push the boulder, Moriarty will do so in such a way that it will be nearly certain to kill Holmes. Watson then makes the decision (event F) to push the boulder, but his skill fails him and the boulder falls on Holmes and kills him (event E).

This example shows that $Q(E: F)$ and $\chi(E: F)$ cannot be identified, since F had a tendency to prevent E and yet caused it. We say that F was a cause of E because there was a chain of events connecting F to E, each of which was strongly caused by the preceding one.

(to be concluded)

Admiralty Research Laboratory
Teddington, Middlesex

COMMUNICATIONAL EPISTEMOLOGY (I) *

MAGOROH MARUYAMA

IN the history of philosophy the philosopher for long asked himself: 'What is the relationship between THE universe and THE human understanding about it?' This question presupposed the existence of THE universe and of THE human understanding.

THE universe was conceived of in three ways:

- (1) As THE material reality outside the human organism, existing independently of the human mind, and perceived through sense organs.
- (2) As THE ideal reality, existing independently of the human mind, and of which the human mind can attain understanding, at least in part, by refined thinking.
- (3) As a construction of the human mind which is the same for all human beings, though it might be different for other beings.

Those who conceived of the universe in the first way, as THE material reality, can be subdivided into:

- (1a) Those who interpreted the natural law as independent of the human mind.
- (1b) Those who interpreted it as a construction of the human mind.
- (1c) Those who identified the human mind with God, with the natural law, with nature, or, as in mysticism, the human mind with God, natural law and nature.

THE universe of (1a), (1c) and (2) did not depend on the existence of THE human mind. But the THE-ness of the universe in (1b) and (3) depended on the THE-ness of the human understanding. Belief in the possibility and success of interpersonal communication was also dependent on belief in the uniqueness and sameness of the human understanding, except for the subjective qualities of sensations and emotions. This belief was strengthened by the fact that philosophy dealt with conscious and verbally communicable realities¹ and by

* Received 3.xi.59. The writer is indebted to a referee for numerous stylistic improvements.

¹ M. Maruyama, 'Communicable and Incommunicable Realities', *The British Journal for the Philosophy of Science*, 1959, 10, 50-54

assuming that the encoding of thoughts into language was the same process for all persons and that encoding was the inverse operation of the decoding or interpreting of the other persons' statements. In fact, thoughts were often identified with the language in which they were expressed. Any unsuccessful communication was attributed to mistakes in THE logical process or to discrepancies in the definitions of words. Further, the assumption of the existence of concepts as ideal universals reduced discrepancies in definitions to discrepancies in the coverage by words of the concepts and subconcepts which, in their totality, were assumed to be the same for all persons and all languages.

The sameness of the human understanding, its isomorphism and homogeneity as between persons, was not even questioned or, at least when found, individual differences were regarded as 'subjective' and 'psychological'; they were omitted from philosophical discussions as unworthy of consideration. This was the mechanism by which differences in human understanding were eliminated, enabling epistemology to be founded on the 'universality', 'integrity' and 'absoluteness' of human understanding, thereby making itself universal, integral, and absolute also.

These restrictive assumptions and the classical epistemology based on them are here discarded in favour of a 'communicational epistemology' which recognises human differences. To recapitulate, the restrictive assumptions discarded are:

- (1) That human understanding is the same in all persons (or that epistemology is restricted to the study of those parts of human understanding that are the same for all persons).
- (2) That logic is the same for all persons (or that 'reasoning pattern' is the same for all persons, if logic is to be reserved for Aristotelian deductive logic).
- (3) That concepts are ideal universals identical for all persons.
- (4) That epistemology is restricted to the study of conscious and verbally communicable realities.
- (5) That the encoding of thoughts into statements, actions, etc. is the same for all persons.
- (6) That the decoding of statements, actions, etc., is the same for all persons.
- (7) That decoding is the inverse function of encoding.

Hence communicational epistemology is based on:

COMMUNICATIONAL EPISTEMOLOGY

- (1) The existence of various patterns of human understanding.
- (2) The existence of various reasoning patterns.
- (3) The variability of concepts as between persons.
- (4) The existence of unconscious communication as an epistemological subject. The existence of nonverbal communication as an epistemological subject. The existence of incommunicable realities.
- (5) The existence of persons with different encoding functions.
- (6) The existence of persons with different decoding functions.
- (7) The existence of persons whose decoding functions are not the inverse of their encoding functions.

If these seven conditions are not affirmed, then the inauguration of communicational epistemology is not justified. If, on the other hand, any of the conditions are affirmed, inauguration of communicational epistemology is a necessity. But recent developments in cultural anthropology and psychiatry affirm the seven conditions.¹

Three aspects of classical epistemology have to be modified by acceptance of the seven conditions.

They are:

- (1) The relationship between the universe and the human mind.
- (2) The relationship between the different understandings of the universe by different persons.
- (3) The relationship between the understandings by different persons of each other, or the problem of interpersonal communication.

In this present paper only the third aspect is discussed.

The study of interpersonal communication may be divided into three stages:

- (1) Study of the confusions and misunderstandings resulting from contacts between persons with different patterns of thinking, and different encoding and decoding functions.
- (2) Study of methods of detecting misunderstandings, modification of encoding and decoding functions, modification of thinking patterns, criteria for correctness of interpretation, successive approximation and its limits of convergence.
- (3) Study of the limitations of interpretation and of interpersonal communication.

¹ M. Maruyama, *Epistemology of Intercultural Understanding: A Study in Behavioral and Communicational Epistemology; Rediscovery of Philosophy as an Open Meta-Science of Interdisciplinary Cross-Induction*, and its Supplement I, 1958, 1959. Mimeographed

But the question may be asked: 'Are we not now leaving the domain of philosophy and digressing into the domain of psychology?' The answer is 'No'. The easiest way of proving this is to show that communicational epistemology can be treated, though not necessarily, in the most non-psychological way, i.e. by means of relational algebra, formal operations, etc., within the field of formal logic. Each of the three stages of study given above can be treated formally and rigorously¹ But in this paper, to introduce the subject, only simplified illustrative examples will be given.

I A simplified model of misunderstandings due to differences in thinking patterns, encoding and decoding functions

A typical thinking process in Aristotelian logic is to draw a conclusion from a major and a minor premise, using the law of deduction. A typical operation of an electronic computer (a thinking machine) is to calculate the result of applying to the stored and the newly given information the programmed instructions given in advance. The typical function of a decision element in an electronic computer is to combine two inputs to yield one output determined by the wiring of the element—the internal wiring of the element is called the 'logic' of the element. Mathematically, all these three processes, one by a logician, one by a computer, one by a decision element, can be classified as binary functions of the form $f(x,y) = z$, where f is the law of deduction, programming, internal wiring, etc.

x, y are major and minor premises, stored and input information, two inputs, etc., and z is the conclusion, result of the calculation, the output, etc. The function f is often called a binary operator.² Our model of thinking patterns is also a binary function. The process of our model is to produce an internal state (of thinking, feeling, etc.) out of the previous internal state (residual state) and the input, using the operator inherent in the model. The input is either the result of one's own thinking and action, or one's decoding of other persons' statements and behaviour. The residual state and the input are combined and produce a new internal state. The combinations and products can be tabulated in a form similar to that of a multipli-

¹ M. Maruyama, Op. cit.

² The word 'binary' is used here in the sense that there are two independent variables or two inputs. This should not be confused with the 'binary' of 'binary digit', or 'bit' in short, used in the electronic computer. The latter refers to a digital system based on the number 2.

COMMUNICATIONAL EPISTEMOLOGY

cation table. This table is the 'thinking pattern', but it may also be called the 'logic', 'algebra', 'internal process function', 'operator', etc. For example, a person A has the residual state (in this case knowledge) 'B has money' and A's interpretation of B's telephone call is 'B is not thinking of inviting A to a concert'; then A's next internal state (thought) is 'B is angry with A', or 'B is impolite', etc., depending on A's thinking pattern (objective, suspicious, etc.), and on the prevailing customs in A's, but not necessarily B's society, etc. Another example: A's residual state is 'My father is a respectable person', and then A suddenly phantasies himself killing his father (an input from his own thinking). The result may be a feeling of guilt and reinforced submission, with the thought 'My father is an impeccable man'; or it may be a realisation that A does not like his father, with the thought 'My father is full of defects', depending on A's personality structure. Thus, the 'thinking pattern' or 'logic' depends on a person's personality as well as on the culture in which he was brought up.

We can construct a model of the thinking pattern of an intellectual process, but the 'multiplication table' of intellectual processes becomes rather complex. On the other hand, if we take attitude or desire, for example, as the internal state, it is possible to construct a simpler model to serve as an illustration. This is not to imply that the models are valid only for emotional and psychological states; it is emphasised that the method of constructing the model can be applied also to intellectual processes.

For each combination of the residual state and the input, the new state may not be determined uniquely but may have several alternatives with a probability distribution. For the sake of simplicity, let us assume that the new state is uniquely determined. Different combinations of the residual state and the input may result in a same new state—this possibility is not excluded. An input which is the result of one's own thought or action is called the *return*. An input which is the result of the decoding of another person's statement or action is called the *reception*. For example, a person A codes his contempt for another person B in a form of smiling and B decodes the smile as a sign of A's friendliness towards B. Then 'A is friendly to me' is B's reception. The transformation from A's smile to 'A is friendly to me' is B's decoding. For the sake of simplicity let us assume that the return and the reception have the same effect on the residual state, i.e. as far as the algebra is concerned, it does not matter whether 'A is friendly to me' is a result of B seeing A's smile or is simply a thought which occurred to B without A's presence.

Let us further assume that, in reception, any decoding is transferred to one's own state as input; for example, that my reception of 'You are friendly' makes me friendly. This is called the *assumption of simplicity*. If the state caused by return is again the same state as input, the return is called *regenerative*. For example, the state of desiring to see films making one go to a movie has a regenerative return if this action results in the desire to see more films.

Let us further assume that there is only a finite number of internal states, inputs, and resulting states in any person, and that the states are discrete, i.e. no continuous change of shading from one state to another exists. For example, suppose that the person A has three possible states *c* (contemptuous), *d* (desiring to dance), and *h* (hungry), and that he decodes other person's behaviour in terms of these three states. Let his thinking pattern of 'algebra' be as follows:

residual state: *c* *d* *h*

input	<i>c</i> : <i>c</i> <i>c</i> <i>h</i>
	<i>d</i> : <i>c</i> <i>d</i> <i>c</i>
	<i>h</i> : <i>c</i> <i>d</i> <i>h</i>

This table is called the *state algebra* or *stal* of A. The combination of the input and the residual state is expressed in the form of a product, the left-hand factor representing the input and the right-hand factor the residual state. Thus, $dh = c$ means that the combination of the residual state *hungry* and the input *desiring to dance* results in the new state *contemptuous*. A state algebra has to satisfy the following condition: If *S* be the set of states, then for any arbitrary elements *a* and *b* of *S*, ab is uniquely determined and is again in *S*.

It should be noted that neither commutativity nor associativity is assumed, i.e. that neither $ab = ba$ nor $(ab)c = a(bc)$. Any subset of a *stal* is called a *substal* if it satisfies the conditions of a *stal*.

So far we have formalised the thinking pattern of only one person. The next step is to examine what happens when two persons with different thinking patterns and different encoding and decoding functions try to communicate with each other. For this purpose let us formulate how A codes his internal state into behaviour. Let *c* be coded into a smile, *d* into talking much, and *h* into putting down a cigarette. Now we have to specify how A decodes other persons' behaviour. It would be natural to assume that since A smiles when A is contemptuous, A interprets anyone else's smile as a sign of contempt: if this were so, A's decoding function would be the inverse of his

COMMUNICATIONAL EPISTEMOLOGY

encoding function. But a person's encoding and decoding functions are not necessarily inverse. An Eskimo living in New York may behave as an Eskimo but may know how to decode an American's behaviour correctly. Or a slave may behave in a slavish way and yet know how to decode his master's behaviour. In these cases, the decoding function is not the inverse of the encoding function. But again for the sake of simplicity, let us assume that the decoding of a person is the inverse of his encoding.

If, for every person within an interacting group, the decoding function is the inverse of the encoding function and all returns have the same effects as the receptions, we call the interaction between the persons *prime*. We will also make the further simplifying assumption that the results of encoding and decoding are uniquely determined, i.e. that the encoding and decoding are *univocal*. When the states are idempotent (i.e. if, when s is a state, $ss = s$) then the return has no effect on the residual state and can therefore be neglected.

We now introduce another person B. Let him have three states: f (friendly), n (do not believe you) and d (desiring to dance). Let his stal be:

	f	n	d
f :	f	n	d
n :	f	n	d
d :	d	f	d

And let B code f into a smile, n into talking much, and d into putting down a cigarette. We assume that B's decoding is the inverse of his coding. Furthermore, to make the situation more realistic, let us change our first person A into a lady.

We then obtain the following links between the two persons through their behaviour:

A		Behaviour	B			
<i>c</i>	<i>d</i>	<i>h</i>		<i>f</i> <i>n</i> <i>d</i>		
<i>c</i> : <i>c</i>	<i>c</i>	<i>h</i>	<i>c</i> (contempt)	smile	<i>f</i> (friendly)	<i>f</i> : <i>f</i> <i>n</i> <i>d</i>
<i>d</i> : <i>c</i>	<i>d</i>	<i>c</i>	<i>d</i> (desiring to dance)	talks much	<i>n</i> (do not believe you)	<i>n</i> : <i>f</i> <i>n</i> <i>d</i>
<i>h</i> : <i>c</i>	<i>d</i>	<i>h</i>	<i>h</i> (hungry)	puts down cigarette	<i>d</i> (desiring to dance)	<i>d</i> : <i>d</i> <i>f</i> <i>d</i>

As all states are idempotents, the returns may be neglected. This diagram with the stals of A and B flanking their behaviour pattern is a very simple model of the interaction between two persons. A model of interpersonal interaction thus schematised is called an 'algebra of interpersonal interaction' or *intal*. When an *intal* is given and the initial states of the persons are specified, we can compute the sequence of interactions between the persons.

For example, let us assume that A's initial state is *h*, that B's initial state is *n* and that A expresses herself first. The course of interaction develops as follows: A puts down her cigarette, an action which B interprets as indicating that A desires to dance, i.e. his reception is *d*. The combination of reception *d* with B's initial state *n* produces *f*, a state of friendliness which B expresses with a smile. This is received by A as *c*, an expression of contempt. The reception *c* combined with A's residual state *h* produces the state *h*: A puts down her cigarette. This gives reception *d* on B. But this time B's residual state is *f*, and the combination of reception *d* and residual state *f* is the new state *d*; B is in the state of desiring to dance. He expresses this state by putting down his cigarette, which makes reception *h* on A. The result, the combination of reception *h* with residual state *h*, is again *h*. From this point the process repeats, A and B putting down their cigarettes alternately, while A is always hungry and thinks that B is hungry, while B is always anxious to dance and thinks that A is also anxious to dance. This kind of misunderstanding is typical of intercultural contact situations.

But if the initial states of the two persons are different from those in the above example, the sequence of interactions may take a different course and may end with a different final behaviour. If, for example, A starts with *c* and B with *d*, and if A expresses herself first, then, after some computation the final state reached is that A is always contemptuous and smiles believing that B is hungry while B is always anxious to dance and puts down his cigarette, thinking that A is friendly. This also is a typical pattern of intercultural misunderstanding.

Our example ended with situations in which an equilibrium is established, each person remaining unchanged. But there are *intals* which never produce equilibrium regardless of the choice of the initial states of the two persons and regardless of who expresses himself first. Suppose we change the stals of the preceding *intals* as follows:

COMMUNICATIONAL EPISTEMOLOGY

A	<i>c</i>	<i>d</i>	<i>h</i>	B	<i>f</i>	<i>n</i>	<i>d</i>
	<hr/>				<hr/>		
	<i>c:</i>	<i>d</i>	<i>h</i>		<i>f:</i>	<i>d</i>	<i>d</i>
		<i>h</i>	<i>c</i>			<i>n</i>	<i>d</i>
	<i>d:</i>	<i>h</i>	<i>h</i>			<i>f</i>	<i>f</i>
		<i>d</i>	<i>c</i>			<i>n</i>	<i>f</i>
	<i>h:</i>	<i>d</i>	<i>c</i>			<i>n</i>	<i>n</i>

and assume that there is no return. Then we find that both A and B ultimately reach a cycle of three states. The pattern of interactions and misunderstandings as determined by the structure of the intal (which is composed of the stals and the encoding and decoding functions of the persons involved) can be elaborated. The procedure, as we have seen, is a strictly logical and formal process. Though prosaic interpretations have been used in the illustrative examples, the interactions can be formalised. It is also possible to construct electronic robots with given stals, encoding, and decoding functions, and to make them play out interpersonal misunderstandings. The wiring diagram of such robots is available.¹

We have seen how the first stage of our inquiry into communicational epistemology—misunderstandings due to differences in thinking patterns, encoding and decoding functions—can be formalised. We next present a formal treatment of the second and third stages of our inquiry—correction of misunderstandings and the limitations of interpersonal understandings. Empirical examples of misunderstandings in interpersonal and intercultural communications will then be given.

(to be continued)

¹ M. Maruyama, 'The Circuitry of Electronic Robots to play out Interpersonal Misunderstandings.' 1957. Mimeographed

DISCUSSIONS

A SECOND NOTE ON THE PRINCIPLE OF MINIMUM ASSUMPTION

ALTHOUGH I do not believe that Professor Kapp has satisfactorily met the points I have raised against his thesis¹ I shall not go over the same arguments again. It is better left to the judgment of the interested reader to decide whether the proposed principle of minimum assumption has held the ground against these points.

There was, however, one point to which Professor Kapp has made no reference at all in his reply and which by itself would suffice to show the inadequacy of his principle. I am referring to the last objection raised in my note that Kapp's principle lacks constraining conditions and therefore, like all minimal (or maximal) principles which lack constraining conditions, is indeterminate. I now realise that this was a rather curt statement and its significance none too clear. I should like, therefore, to elaborate slightly upon it.

To justify this further imposition on the patience of the readers of this journal I venture to suggest that the point I am about to make may be of more general interest. It provides a reason why all assertions to the effect that nature is governed by maximum simplicity are indefensible.

Let us consider Fermat's principle of least time, from which it is said one can determine the path of a light ray traversing different media and suffering successive refraction. Let us state it as follows: 'The path travelled by light going from one point to another, traversing different media, is that which requires least time.' Stated just like this, the path of light would be indeterminate. The assumption which would conform best with the principle would be that the velocity of light was always infinite. Hence, whatever the shape of the path followed by light, it will travel from any point A to point B instantaneously. But of course Fermat's principle has to be applied in conjunction with the constraining conditions, which are: the velocity of light is finite and is uniquely determined by the nature of the medium traversed. The time spent in transit is now wholly dependent on the shape of the path. Given the velocity of light in the various media we can construct a unique course which is to be followed so as to minimise the duration of the journey.

Let us now consider again Kapp's principle: 'In Physics the minimum assumption always constitutes the true generalisation.' Clearly, the state of affairs which would conform best to this principle would be that in which there were no forces, no action, neither matter nor motion—a universe in which nothing existed but still void. Professor Kapp's Cosmic Statute Book would then contain the minimum number of entries; that is, no entries at all.

Undoubtedly then, Kapp's principle as it stands is incomplete. Lacking appropriate constraining conditions, it is not a principle of minimum assumption but a principle of no assumption, which is continually violated by the presence of any phenomena.

The only doubt one may still have is: could there be a remedy for this situation? And the answer would have to be: by stipulating that nature has certain fixed

¹ This *Journal*, 1960, 10, 55-62

REPLY TO NOTE BY G. SCHLESINGER

objectives and that Ockam's razor is only to be applied to the way in which these objectives are achieved but not to the objectives themselves. We would have to distinguish between 'ends' and 'means' in nature. The 'ends' of nature are set but the 'means' must comply with the principle of minimum assumption.

This basic distinction has to be made whenever it is claimed (irrespective of the particular form in which it is claimed) that in nature maximum simplicity obtains. It is always understood that maximum simplicity applies not to all phenomena but to those which are instrumental in bringing about the objectives of nature.

Thus, for example, Galileo's dictum, 'Nature does not do by many that which can be done by a few'. It implies that there are certain tasks to be 'done' by nature and these tasks are executed with maximum efficiency and economy. One could, however, not speak of efficiency and economy if there were no specific tasks to accomplish.

The difficulty then is obvious. One cannot seriously defend a classification of phenomena into 'ends' and 'means'. Hence, already on this account the principle of minimum assumption is untenable.

Finally, it should be clear that this difficulty does not defeat all the attempts to formulate a principle of simplicity. It obviously does not affect the view in which the principle is not a feature to which the universe conforms but a rule for the scientist to conform with. 'Given a certain phenomenon, choose the hypothesis of maximum simplicity.' The constraint here is set by the observer. The 'ends' are the explananda and the 'means', which have to be kept at a minimum complexity, are the explanations.

But even a claim that nature itself is simple can escape the difficulty provided no maximum simplicity is stipulated. There would be no *a priori* objection on the above grounds to a thesis that all the relationships in nature are simple enough to be representable by a mathematical expression not more complex than F —where F is a given functional relationship. But no version of the claim that this universe of ours is the simplest possible can be defended.

G. SCHLESINGER

REPLY TO NOTE BY G. SCHLESINGER

For the convenience of readers who do not have easy access to the correspondence between Mr Schlesinger and myself in the *Journal* of May 1960, I shall begin by repeating some of the implications of the Principle of Minimum Assumption in all their uncompromising starkness. If this principle is true, all general statements in physics are either tautologies or reducible; they can be inferred from wider generalisations. In this sense they can all be explained. The laws of physics are of a different kind from those laws that demand certain things and prohibit others. There is nothing in physics to which the metaphor of a cosmic statute book applies. Nature cannot be regarded as a legislator. There is no law or principle in physics that makes for specified order in the sense in which the laws of a country make for specified order.

The question with which this correspondence is concerned is whether the Principle of Minimum Assumption is valid. To formulate it with philosophical precision

would require a long and cumbersome dissertation, but this is not necessary. The question can be paraphrased: Is there, or is there not, such a thing (metaphorically speaking) as a cosmic statute book? In my youth I took it for granted, and without giving any thought to the matter, that there must be such a book and that it must contain a great many entries. In middle age I was seriously questioning this. Today I am almost satisfied that there is no such book at all. Mr Schlesinger thinks there probably is, but I am sure he would agree that it cannot contain as many entries as it seemed to in the past.

The disagreement therefore centres around the 'some or none' issue; entries or no entries. But it is important also to note well the point on which we are agreed. This is that the question deserves serious study and calls for disciplined thinking and discussion; I myself would regard it as one of the two or three most basic unanswered questions in the philosophy of science. The trouble is that both possibilities—'some' and 'none'—have unwelcome implications, which explains the unpopularity of the question. It is easier to change the subject or to suggest vaguely that the correct answer may lie somewhere between 'some' and 'none' than to face the difficulties raised by each possibility. I welcome it that Mr Schlesinger is prepared to face the difficulties.

Let me now give two reasons why my own satisfaction with the Principle of Minimum Assumption, although ninety per cent, is not a hundred per cent.

Firstly, there is Mr Schlesinger's own objection that the principle implies a complete absence of restraining conditions. This is true by definition. Mr Schlesinger says that, in the absence of restraining conditions, the material universe would be indeterminate. I agree. But I could also imagine this argument being used against Mr Schlesinger; for the material universe *is* indeterminate. The movement of individual electrons conforms to the uncertainty principle; the moment of time when a given atom of radiation disintegrates is not determined by anything in the existing state of affairs. There are certainly events that occur in the absence of restraining conditions. However an attempt to counter Mr Schlesinger by quoting these would only be to score a cheap debating point. While it is true that individual micro-events are indeterminate it also follows from the laws of probability that the effect of a large collection of events is predictable. If the cosmic statute book contains nothing else it should, according to Mr Schlesinger's argument, contain at least those fractions that define the probabilities of certain classes of events. I do not think that Mr Schlesinger's argument is easy to refute.

Secondly, there are certain statements in physics that appear to be incapable of reduction, of explanation; no one has succeeded in showing that they are implicit in wider generalisations or that they are tautologies. Among them are the conservation laws, a handful of cosmic constants, some quantitative definitions of probability. These suggest a cosmic statute book. My belief that they will one by one disappear, as so many other generalisations have done in the past, is based only on faith and not on facts. This does not allow me to be a hundred per cent satisfied with my own Principle.

But against all this, let me now say why my doubts about the validity of this Principle are really very slight.

Firstly, the last few centuries have seen a most consistent and drastic elimination of the sort of 'laws' that could originally have led to the notion of a cosmic statute

INDIVIDUALISM

book. The number that seem to remain look, by contrast, so meagre, so trivial, so odd, that I find it more rational to believe that even these will be eliminated in due course; that it will eventually be seen that there never were such things as 'Nature's Laws.'

Secondly, there is a pragmatic consideration. To accept that a generalisation in physics implies a cosmic statute book is to treat it as inexplicable; it is to abandon all hope that it may some day be understood and correlated with other generalisations. This consideration suggests that all such general statements in physics should be critically examined with a view to their possible elimination from the cosmic statute book. Whether the Principle of Minimum Assumption is valid or not, physicists do in fact often act as though they believed in it, and their action has often proved fruitful.

Years ago it occurred to me to examine critically two generalisations all too readily treated as though they were entries in our imaginary book. They were: 'The contents of the material universe had to originate at a specific point in time, a finite number of years ago', and 'Every elementary component of the material universe shall continue to exist for an infinite time.' It was my discovery of how much that was puzzling could be explained by denying these generalisations that increased my fifty per cent satisfaction with the Principle of Minimum Assumption to its present ninety per cent.

REGINALD O. KAPP

METHODOLOGICAL AND EPISTEMOLOGICAL INDIVIDUALISM

HAYEK and Popper have something important to say on methodological individualism. I have been reading more recent writings on this subject, and I wish to report that the game's not worth the candle. The more recent writings I have read comprise eight contributions to this *Journal*, various contributions to *Philosophy of Science*, *The British Journal of Sociology*, the *Aristotelian Society Supplementary Volumes*, and *The Journal of Philosophy*, and parts of several books.¹

Two different principles have the attractive designation 'methodological individualism'. Hayek applies the term to a methodological principle, and Popper applies it to an epistemological principle. Hayek's principle is one about the *methods* that should be employed in gathering information and forming theories: our *data* in the social sciences are 'the relations between individual minds which we directly know'.² Popper's principle is a blanket epistemological principle: whatever methods we have used, 'we should never be satisfied by an explanation in terms of so-called "collectives" (states, nations, races, etc.)'.³ Hayek's principle is about how we should start our enquiries, Popper's about how we should finish them: Hayek says that the methodological individualist 'systematically *starts from* the concepts which

¹ Many of the references are given in footnotes to J. W. N. Watkins, 'Historical Explanation in the Social Sciences', this *Journal*, 1957, 8, 104

² F. A. Hayek, *The Counter-Revolution of Science*, Illinois, 1952, p. 57

³ K. R. Popper, *The Open Society and its Enemies*, London, 2nd edn., 1952, Vol. 2, 98

guide individuals in their actions';¹ Popper says that methodological individualism 'rightly insists that the "behaviour" and the "actions" of collectives, such as states or social groups, must be *reduced* to the behaviour and to the actions of human individuals'.² Hayek's principle is synthetic, Popper's analytical: Hayek says that the social sciences 'do not deal with "given" wholes but their task is to *constitute* these wholes by *constructing* models from the familiar elements';³ Popper says that 'institutions (and traditions) must be *analysed* in individualistic terms'.⁴ It is not surprising that there should be this series of related differences between Hayek and Popper, for Hayek is interested only in how theories are formed and says nothing about testing theories, and Popper says:

The question, 'How did you first *find* your theory?' relates, as it were, to an entirely private matter, as opposed to the question, 'How did you *test* your theory?' which alone is scientifically relevant.⁵

I have read eight articles and notes by Watkins on methodological individualism, five of them in this *Journal*. He does not distinguish between methodological individualism properly so called and what I have called epistemological individualism. He writes as though Hayek and Popper mean the same thing by 'methodological individualism'. He advocates a 'methodological individualism' which is a jumble of Hayek's principle and Popper's principle. The jumbling process has distorted both of these principles without producing any new principle.

Hayek says that 'in the social sciences our data or "facts" are themselves ideas or concepts'⁶ in the minds of individuals. Even statistics, which one might expect an economist to regard as furnishing important data, are (unless they are 'concerned with the attributes of individuals')⁷ quite irrelevant to social theory, since they do no more than 'provide us with the data to which our theoretical generalizations must be applied to be of any practical use'.⁸ This may be sound or unsound, but it is at least unqualified methodological individualism.

Watkins distorts Hayek's methodological individualism when he says:

The assertion that knowledge of social phenomena can only be derived from knowledge about individuals requires one qualification. For there are certain overt features which can be established without knowledge of psychological facts, such as the level of prices, or the death-rate (but *not* the suicide-rate).⁹

He gives a circular definition of 'overt feature' as 'something which can be ascertained without referring to people's dispositions, etc.'¹⁰ In other words, Hayek's universal principle applies to all social phenomena except the ones it does not apply to.

Hereafter in deference to Popper I shall drop the term 'epistemological individualism'. Popper says that methodological individualism is the doctrine that 'we must try to understand all collective phenomena as due to the actions, interactions,

¹ Op. cit. p. 38 (my italics) ² Op. cit. p. 91 (my italics)

³ Op. cit. p. 56 (my italics of second word italicised)

⁴ Op. cit. p. 324 (my italics)

⁵ K. R. Popper, *The Poverty of Historicism*, London, 1957, p. 135

⁶ Op. cit. p. 36

⁷ Op. cit. p. 61

⁸ Op. cit. p. 63

⁹ J. W. N. Watkins, 'Ideal Types and Historical Explanation', this *Journal* 1952, 3, 28 (Watkins's italics)

¹⁰ Ibid.

INDIVIDUALISM

aims, hopes, and thoughts of individual men, and as due to traditions created and preserved by individual men'.¹ At first sight it is puzzling that Popper should superadd this requirement to what he so often says is his sole requirement, testability. Thus it is not easy to reconcile his methodological individualism with his statement that 'the question, "How did you *test* your theory?" alone is scientifically relevant'. I believe that to solve this puzzle we have to examine Popper's views on the purpose of social science. This purpose is twofold: first, to form 'sociological laws or hypotheses which are analagous to the laws or hypotheses of the natural sciences';² second to give 'the explanation of some regularity or law'.³ The solution to the puzzle is that in a law all that matters is testability, but an explanation must comply with the principle of methodological individualism. Laws do not need to comply with this principle. Most of Popper's examples of sociological laws⁴ contravene the principle of methodological individualism, for instance 'You cannot have a centrally planned society with a price system that fulfils the main functions of competitive prices' and 'You cannot have full employment without inflation'.

Watkins distorts Popper's principle (and shows that he forgets the purpose of science) when he down-grades Popper's 'laws' to 'unfinished or half-way [*sic*] explanations':

According to this principle, the ultimate constituents of the social world are individual people who act more or less appropriately in the light of their dispositions and understanding of their situation. Every complex social situation, institution, or event is the result of a particular configuration of individuals, their dispositions, situations, beliefs, and physical resources and environment. There may be unfinished or half-way explanations of large-scale social phenomena (say, inflation) in terms of other large-scale phenomena (say, full employment); but we shall not have arrived at rock-bottom explanations of such large-scale phenomena until we have deduced an account of them from statements about the dispositions, beliefs, resources, and inter-relations of individuals. (The individuals may remain anonymous and only typical dispositions, etc., may be attributed to them.)⁵

Watkins gives two illustrations of theories that his principle would prohibit. To be satisfactory, these would have to be illustrations of either (a) explanations that do not comply with the principle or (b) 'half-way' explanations that cannot be reduced to explanations that comply with the principle. They both prove, on examination, to be illustrations of (c) 'half-way' explanations that can be so reduced. As 'half-way' explanations they are laws and hence acceptable in Popper's view; and, even if Popper required laws to be reducible to individualistic statements (which he does not do), these explanations would still be acceptable to him since they are so reducible. The first illustration is:

An example of such a superhuman, sociological factor is the alleged long-term cyclical wave in economic life which is supposed to be self-propelling, uncontrollable, and inexplicable in terms of human activity, but in terms of the fluctuations of which such large-scale phenomena as wars, revolutions, and mass emigration, and such psychological factors as scientific and technological inventiveness can, it is claimed, be explained and predicted.⁶

¹ Popper, *The Poverty of Historicism*, pp. 157-158

² *Ibid.* p. 62

³ *Ibid.* p. 122

⁴ *Ibid.* pp. 62-63

⁵ This *Journal*, 1957, 8, 105-106

⁶ *Ibid.* pp. 106-107

This illustration is unsatisfactory unless Watkins can produce an economist who purports to propound a theory that is inexplicable in terms of human activity. This he does not do. It is worth noting that in a footnote to a very similar passage in his 1952 article Watkins says that he wrote the passage with the Russian economist Kondratieff in mind. As he does not repeat the footnote in his later article, he possibly no longer believes that Kondratieff illustrates his point. In the article referred to by Watkins, Kondratieff does not say that the long-term cyclical wave is 'self-propelling, uncontrollable, and inexplicable in terms of human activity'. What Kondratieff does say is:

In asserting the existence of long waves and in denying that they arise out of random causes, we are also of the opinion that the long waves arise out of causes which are inherent in the essence of the capitalistic economy. This naturally leads to the question as to the nature of these causes. We are fully aware of the difficulty and great importance of this question; but in the preceding sketch we had no intention of laying the foundations for an appropriate theory of long waves.¹

The second illustration is: 'Marx, for instance, professed to believe that feudal ideas and *bourgeois* ideas are more or less literally generated by the water-mill and the steam-engine'.² But this is a travesty. What Marx says is fully in accord with methodological individualism:

M. Proudhon the economist understands very well that men make cloth, linen or silk materials in definite relations of production. But what he has not understood is that these definite social relations are just as much produced by men as linen, flax, etc. Social relations are closely bound up with productive forces. In acquiring new productive forces men change their mode of production; and in changing their mode of production, in changing the way of earning their living, they change all their social relations. The hand-mill gives you society with the feudal lord; the steam-mill, society with the industrial capitalist.³

There are further confusions in what Watkins has to say. Popper clearly distinguishes methodological individualism from methodological psychologism:

The mistake of psychologism is its presumption that this methodological individualism in the field of social science implies the programme of reducing all social phenomena and all social regularities to psychological phenomena and psychological laws.⁴

Watkins advocates a psychologicistic methodological individualism in his 1952 article in this *Journal*, and an anti-psychologicistic methodological individualism in his 1957 article in this *Journal*. In 1952 he gave a classic statement of methodological psychologism:

From this truism I infer the methodological principle which underlies this paper, namely, that the social scientist can continue searching for explanations of a social phenomenon until he has reduced it to psychological terms.⁵

¹ Kondratieff, in source cited by Watkins, pp. 41-42

² This *Journal*, 1957, 8, 111

³ Marx, *The Poverty of Philosophy*, London, Martin Lawrence, not dated, p. 92

⁴ Popper, *The Open Society*, Vol. 2, p. 98

⁵ This *Journal*, 1952, 3, 28-29

INDIVIDUALISM

And in 1957, without any apparent awareness that he had changed his views, he said:

Another [misunderstanding] of methodological individualism is that it has been confused with a narrow species of itself (Popper calls it 'psychologism'). . . .¹

There is of course no reason why a writer's views should not change over the years, but when they change he should not confuse his readers by leading them to suppose that his new views are re-statements of his old views.

In the same 1957 article there appears an economic digression which Watkins might have supposed runs counter to any version of methodological individualism:

. . . it is very doubtful whether an economist can ever *show* that an economic system containing negative feed-back will be stable. For negative feed-back may produce either a tendency towards equilibrium, or increasing oscillations, according to the numerical values of the parameters of the system. But numerical values are just what economic measurements, which are usually ordinal rather than cardinal, seldom yield. The belief that a system which contains negative feed-back, but whose variables cannot be described quantitatively, is stable may be based on faith or experience, but it cannot be shown mathematically.²

Watkins does not appear to have replied to what I consider the most cogent of the criticisms of methodological individualism, Maurice Mandelbaum's 'Societal Laws', which appeared in this *Journal* in 1957.³ This is odd, for Watkins has replied to most of the criticisms levelled against his versions of methodological individualism, and has replied to a criticism which appeared in this *Journal* six months later than Mandelbaum's article. Mandelbaum gives a clear-headed classification of sociological laws into four classes and shows point by point that his fourth class (which he calls 'abstractive-functional societal laws'), though prohibited by the principle of methodological individualism as it is stated by Watkins, are free from the defects that Watkins attributes to the theories his principle prohibits.

It is true that Mandelbaum is non-committal on whether such laws exist:

Whether such laws have been found, or whether we have reason to believe that they may be found, is not the question which I have proposed for this discussion.⁴

But Mandelbaum also says:

Among the examples of attempts to formulate such laws we may cite the following: statements concerning relationships between modes of production and marriage systems; between size of population and political organisation; between forms of economic organisation and political organisation; or, to cite a classic study of Tylor's (which has been amplified and elaborated by Murdock in his *Social Structure*) between certain specific aspects of marriage systems, e.g. rules of residence and, rules of descent.⁵

Mandelbaum's distinction between 'laws' and 'attempts to formulate such laws' is unsatisfactory in terms of Popper's refusal to accept a distinction between hypotheses and laws.⁶ If attempts have been made to formulate laws, then hypotheses have been advanced. But even if Mandelbaum had not in effect stated that hypotheses of his fourth class have in fact been made, Watkins would still need to reply, for the point

¹ This *Journal*, 1957, 8, 111-112

² Ibid. p. 114

³ 'Societal Laws', this *Journal*, 1957, 8, 211

⁴ Ibid. p. 222

⁵ Ibid. p. 221

⁶ For instance, Popper, op. cit., Vol. 2, pp. 260-261

K. J. SCOTT

at issue between Mandelbaum and himself is about a methodological prescription. Watkins says that certain kinds of theory may not legitimately be postulated as they have certain defects, and Mandelbaum says that one class of these theories are free from the defects Watkins attributes to them.

K. J. SCOTT

Victoria University of Wellington

NOTE ON THE TEACHING OF HISTORY AND PHILOSOPHY OF SCIENCE

Department of History and
Philosophy of Science,
University of Aberdeen

Sir,

I SHOULD be obliged if you would correct a statement in the November issue of the *Journal* to the effect that at the University of Aberdeen the balance in the teaching of History and Philosophy of Science is ' Largely in the History of Science '. Though the policy of the Department, in consultation with the Faculties, is to maintain the balance as evenly as possible, it is the case that the philosophical aspects somewhat predominate. Even in the form set out in the article the excess of history is very slight; but what may not have been made clear in my report to Dr Mays is that whereas the weekly philosophy course covers the whole session the history course is only for two terms. Also the evening staff Seminars generally have a philosophical balance.

In conclusion I should like to emphasise that in *all* the courses opportunity is taken to point out the danger of studying either branch of the subject in complete isolation from the other.

Yours faithfully,

WILLIAM P. D. WIGHTMAN

REPLY TO NOTE BY W. P. D. WIGHTMAN

Department of Philosophy,
The University,
Manchester, 13

Sir,

IN reply to Dr Wightman's letter I ought to say that I used the category ' History of Science ' in a fairly wide sense to cover History of Science pure and simple as well as philosophical discussions of science as seen in their historical perspective. I think most of Dr Wightman's discussions are of the latter type. Perhaps the real difficulty arises from my attempt to divorce these two aspects of the subject. It is difficult to do this in practice.

Yours faithfully,

W. MAYS

REVIEWS

WORDS AND THINGS¹

Now that some of the dust stirred up by its first appearance has settled it should be possible to take a somewhat cooler look at Mr Gellner's *Critical Account of Linguistic Philosophy* than was accessible to its first strenuous supporters and opponents. Both its style and its content show it to be a pamphlet rather than a fully-fledged philosophical treatise. It is written in a passionate, hectic and none too orderly way; it is protracted beyond its natural length by a good deal of repetitiousness and its sociological excursions, though entertaining enough as a kind of dinner-table conversation, are only marginally relevant to the book's philosophical intentions. As for its content: Mr Gellner does not exert himself to identify very clearly the precise object of his attack, linguistic philosophy remains to the end a pretty shapeless affair, and his method of condensing basic assumptions out of the philosophical atmosphere does not have much to do with the rules of evidence that ordinarily prevail in this matter. He relies on intuition and hearsay, on revelatory word of mouth, parentheses and stray journalistic lettings-down of hair, rather than on the more solid published endeavours of the objects of his criticism, for a basis on which to erect his account of their fundamental ideas about philosophy. On the rare occasions when he does settle down to the task of examining some concrete piece of philosophical work by a linguistic philosopher the result is sometimes frivolous (cf. p. 217 "uses but not sentences refer" (I have a terrible feeling I may have got the phrasing of this latter idea wrong).) And it is sometimes incompetent (for example, his treatments of the paradigm case argument, of Wittgenstein's rejection of private languages and of Strawson's critique of the theory of descriptions). In a way he has fallen between two stools; a serious critical investigation of a body of philosophical ideas on the one hand, such as Mill's *Examination of Hamilton's Philosophy*, and a satirical, Lytton Strachey-style demolition of an intellectual fashion, such as Fisher's *Our New Religion*, on the other. A shorter and more enjoyable book would have come from concentration on the second task. To have embarked on the first he would need to have done a good deal more work than he has.

The linguistic philosophy he is concerned with is an amalgam of two rather different components: the later philosophy of Wittgenstein and the Austinian philosophy of ordinary language that has dominated post-war Oxford. The adherents of these two bodies of opinion often have a marked

¹ By Ernest Gellner. Gollancz. 1959. 25s.

disregard for one another. But there are some common assumptions that are perhaps most easily defined in terms of the considerable dependence of both groups on the philosophy of G. E. Moore. The very general convictions which Mr Gellner is principally interested in can, I think, most clearly be sorted out under the following three heads (though he does not himself expound linguistic philosophy in quite this way): (a) a doctrine of commonsense naturalism about the world and about language, (b) an insistence on the authority of ordinary language with which is bound up the view that bad, old, metaphysical philosophy consists in subtle and seductive distortions of the normal meanings of words and the view that the proper task of a reformed and disillusioned philosophy is the detection and explanation of these errors, and (c) a kind of flight from theory, an abdication of intellectual responsibility which refuses to propound general philosophical theses on account of the bewilderingly ineffable polymorphism of language.

The merit of this classification is that it conveniently brings out Mr Gellner's main grounds for dissatisfaction with linguistic philosophy. Commonsense naturalism, he claims, is simply a dogma, assumed or insinuated in terms that suggest that no sensible person could doubt it. It fails to deal with the traditional obstacle to naturalism, the epistemology which takes subjective experience to be the only secure basis for knowledge, by simply ignoring it. And in contrast to previous varieties of naturalism it is quite indifferent to the achievements of science, natural or social. This kind of complacent quietism is, however, more conspicuously the offence of school's devotion to ordinary language. All past philosophy is dismissed as confusion, which it is the only possible office of a reformed philosophy to dispel. The outcome is general passive indifference to intellectual creativeness and enterprise and an insincere pretension to neutrality. Succumbing to the authority of ordinary speech the linguistic philosopher neglects his essential duty, that of critically investigating the validity of our inherited stock of ideas and beliefs. In his flight from theory, finally, the linguistic philosopher abandons the search for unity, order, and coherence, which is an indispensable requirement for any serious intellectual undertaking. Such a position underwrites evasiveness and esotericism, and replaces explicit argument by furtive insinuation. In short, dogmatic commonsense and a timid and scholastic conceptual conservatism disdain any thinking that is creative or genuinely critical. They induce the facile acceptance of established ideas and beliefs. The repudiation of theory leads to evasiveness and systematic intellectual dishonesty, to an anti-intellectual obsession with trivial detail and, in the end, to simple boredom.

The first thing to be noticed is that these three doctrines are not very closely logically bound up with one another. A respect for ordinary language naturally presupposes an acceptance of commonsense naturalism

but the converse relation does not hold and both of these views are quite independent of the contention that there can be no general philosophical propositions. A single example is enough to make the point, that of Professor Ayer who has no hostility to conceptual innovation and is certainly no opponent of general theories in philosophy and yet who says 'The philosopher has no right to despise the beliefs of commonsense. If he does so, he merely displays his ignorance of the true purpose of his enquiries.' A difficulty here is that the acceptance of commonsense is an exceedingly vague commitment. Mr Gellner's versions of it—'The world is what it is', 'The world is what it seems'—do nothing to remove the vagueness. Is it a disposition to reject any philosophical doctrine that *conflicts* with commonsense or, more stringently, that merely *adds to* it, by, for example, countenancing the existence of sense-data? And just how far does the authority of commonsense extend? If it holds that men can often rightly be blamed for what they do is it also committed to the occurrence of uncaused acts of will; if it maintains that we really perceive independent material things in space does it also contend, with Professor Price's splendid Naive Realist, that visual sense-data are literally identical with the surfaces of material things? While commonsense naturalism about the world in general is vague, and conspicuously left in that condition by Mr Gellner, commonsense naturalism about language, the view that it is a rule-governed behavioural activity in the common world, seems just obvious. No one could seriously deny that it is at least this. It is to go further, though, to say with Wittgenstein, that it cannot conceivably be anything but this. But that language is a social art—the phrase is Professor Quine's—is hardly a belief peculiar to linguistic philosophers.

Mr Gellner's main objection to commonsense naturalism seems to be that it is simply a dogma, advanced, not on logical grounds, but in an exclusively rhetorical fashion with the aid of scornful utterance and raised eyebrows. This is surely incorrect. G. E. Moore offered a large number of arguments in its defence, conveniently put together in the second chapter of Mr A. R. White's book on him. Professor Quine, amongst others, has made use of Neurath's famous image about repairing our conceptual boat with its own timbers. Others again have preferred to augment the natural authority of commonsense by attempting to refute the traditional sceptical arguments against it, the Cartesian epistemology that Mr Gellner holds linguistic philosophers to have simply ignored. It is understandable that, in trying to arrive at the philosophical opinions of Professor Austin, he should have relied on information about unpublished statements. But in that case it is odd that he should have neglected Austin's famous lectures on 'Sense and Sensibilia', delivered on a number of occasions in post-war Oxford. As to the doctrine of the necessary publicity of language, this was elaborately argued for by Wittgenstein and has been elaborately

examined since the publication of *Philosophical Investigations*; it has been criticised by Mr Strawson and Professor Ayer, defended by Professor Malcolm and Mr Rhees. Mr Gellner's own account of Wittgenstein's reasons for his view is perfunctory in the extreme. In short, as far as commonsense naturalism is concerned, Mr Gellner fails to identify the theory to which he is opposed, is mistaken in claiming that it has been simply taken for granted as a dogma of the school and fails seriously to consider, let alone refute, the arguments on which it rests. Above all his objections are deprived of force by his failure to indicate what alternative he has in mind. (This is, indeed, a prevailing feature of his book.) A kind of Lockean scientism is occasionally hinted at, the view that commonsense is undercut by the discoveries of natural science, as in his attempt to reconstitute Eddington's famous two tables, but it is never explicitly argued for.

Belief in the authority of ordinary language is much more central to linguistic philosophy than commonsense naturalism. Mr Gellner picks out four arguments as pillars of linguistic philosophy and holds the belief in the primacy of ordinary language to rest on three of them: the argument from paradigm cases, the argument about the necessary applicability of each of any pair of contrasted terms and what he calls the generalised form of the naturalistic fallacy which says that actual use is correct use. His counter-arguments are not very impressive. To show that the paradigm-case argument is 'immensely silly' he refers to the distinction between connotation and denotation and claims that it confuses the two. But this is much too simple. The paradigm-case argument is, of course, a variant of a traditional empiricist line of thought: the theory that complex ideas must be analysable into simple ideas being turned into the view that we can only discover what a word connotes or means by reference to the things denoted by it or by the other words with whose connotation its own is identified. On this interpretation the scope of the argument is limited to ostensive words. It may be difficult to decide just which words are ostensive. It may be argued that there are no strictly ostensive words at all. But, as philosophical arguments go, it does not seem notably silly. As for the identification of actual and correct use, Mr Gellner is cheerfully confident that its hash can be adequately settled by a reference to the naturalistic fallacy. But 'correct use' here does not mean 'ideal use' or anything like that; it means 'use that is correct in terms of the established and prevailing rules of language'. Where else are we to look for these rules but in the concrete detail of our linguistic practice? This may seem a trivial point, for have not linguistic philosophers gone on to say that actual, correct, established use is ideal, that it cannot be improved on? I think Mr Gellner is right to suppose that they have sometimes spoken as if they did believe this but in fact no such sweeping contention is required to justify their procedure or would be generally accepted by them. What they would say is that the first task of philosophy

is to get a clear understanding of the conceptual system that we actually have before embarking on reforms. Wittgenstein's more severe view was that the acquisition of such an understanding was the *sole* purpose of philosophical inquiry. But neither view rules out linguistic innovation in general, they only contend that it is a secondary, consequential business for philosophers or, more rigidly, not their business at all. One does not become Canute by failing to assist the incoming tide. There is, anyway, a certain artificiality about Mr Gellner's defence of linguistic reform, for the notion of a philosopher as a conceptual innovator is itself a by-product of linguistic philosophy.

Mr Gellner nowhere does justice to the power and originality of what may be reasonably called the central insight of Wittgenstein's later thought (the crystallisation of a suspicion suggested by the work of G. E. Moore): the theory that the main deliverances of the epistemological tradition are marked by a characteristically paradoxical quality which is produced by, and can be explained by the recognition of, the distortion of the common meaning of words occurring in the arguments leading up to them, to which we are led by seductive models and analogies of their working. Now, right or wrong, this is a forceful and fascinating hypothesis. There is nothing evasive about it. It was presented in a clear and unequivocal manner by its inventor. There is a perfectly definite method of verification attached to it, namely the discovery of effective explanations of the paradox-producing distortions, of real cures for philosophical headaches. Even though many who have been much attracted by it would now admit that it has not been very convincingly verified, and that it indicates a contributing factor in philosophical perplexity rather than gives a total explanation of it, they would rightly hold that it has entirely and irreversibly changed the face of philosophy. For it has made it impossible henceforth to take the more counter-intuitive pronouncements of traditional philosophy at their face value as statements of some higher truth about the world than can be communicated by the propositions of science and commonsense. And nothing supports this more effectively than Mr Gellner's contented acceptance of the view that true philosophy is conceptual reform.

Wittgenstein's asceticism about innovation is no longer a matter of theory: for Austin it was a question of establishing a rational order of priorities, for Mr Strawson (in *The Revolution in Philosophy*) it is, as well as this, a matter of personal preference and interest though he is sceptical about philosophical innovation to the extent of seeing the construction of ideal languages as largely the provision of analytical tools for the greater understanding of our actual conceptual system in much the same way as a modern economist might see the relation between classical, Ricardian economic theory and the realities of economic life. An essential point to notice is that conceptual innovation by past philosophy has been generally unconscious

and unintended. When Professor C. I. Lewis says 'no empirical statements are certain' he is presenting what he believes to be an important discovery, not a redefinition of 'certain' as 'completely verified'. It is to concede at least half the linguistic philosopher's claim about ordinary language to recognise as innovation what has hitherto been regarded as discovery.

All the same there is some justice in Mr Gellner's assault on the general attitude towards speculative innovation adopted by linguistic philosophers. The Austinian philosophy of present-day Oxford derives as much from the local school of Cook Wilson and Prichard as it does from Wittgenstein and it does tend to share some of that school's pedantic obsession with the conceptual *status quo*, an outlook which found its most farcial and exemplary expression in H. W. B. Joseph's attempted demolitions of Freud and Einstein. Again his strictures on Austin's evolutionary argument in favour of ordinary language have a good deal of force though he rather excitedly overrates the importance its propounder attached to it. Austin's point was that there was good reason for taking ordinary language seriously and for doubting whether it could be blithely written off as a tissue of inconsistencies and be reformed or reconstructed *a priori*. His argument is a caution against linguistic utopianism. And conservatives are not the only opponents of utopias. Nevertheless the general tendency of ordinary language philosophy has been towards an ultimately uncritical acceptance of the established conceptual order. Mr Gellner's attack on it comes at a time when it is already changing its direction on account of internal developments within linguistic philosophy, above all perhaps a realisation that philosophical problems obstinately refuse to give up the ghost however minute and delicate the methods of ordinary-language analysis that are applied to them. Another factor has been an increasing doubt about the solidity and definiteness of the boundary between the philosophical and the ordinary. Mr Gellner's criticism of the pretension to neutrality which is based on faith in this boundary correspond, in their somewhat boisterous way, to the doubts of many linguistic philosophers.

I shall not spend much time on Mr Gellner's attack on the fourth pillar of linguistic philosophy, the polymorphism that insists on the infinite and ineffable variety of language, and on the flight from theory that it supports since I agree with it. Here again, however, it is necessary to distinguish between Wittgenstein, for whom the impossibility of general philosophical theories was a matter of doctrine, and the Oxford philosophers for whom it is a counsel of prudence. As a doctrine it is plainly inconsistent, not only with Wittgenstein's actual practice, for all his bewildering attempts to stick to pure description, but also with itself. Its powerful hold over Wittgenstein means simply that the interpretation of his later works is made laborious in a new but not very interesting way. There are plenty of instances of the extraction of general theses from the later philosophy of Wittgenstein

and the writings of such followers as Professor Malcolm reveal in practice an admission that some of them would be prepared to concede in principle: that the doctrine of philosophical ineffability is an aberration inherited from the *Tractatus*. Its source would seem to lie in the consideration that philosophy, understood as an inquiry into the relation between language and the world, can only put itself outside the language it is concerned with by saying nothing. But this is a *reductio ad absurdum* not a conclusion to rest content in. In order to discuss language you have to use it, a state of affairs in which it is necessary to be on your guard against begging the question. We are back in Neurath's boat, language cannot be investigated in a condition of perfect linguistic innocence. The best corrective for this conceptual variety of original sin is imagination, the envisaging of languages deeply and substantially different from our own. As a counsel of prudence the polymorphic avoidance of theory is unattractive and Mr Gellner's thundering indictments of evasiveness have some foundation. But evasiveness is neither a monopoly of linguistic philosophers nor is it generally characteristic of them. The vanity of thinkers will always look for a refuge from criticism, as much in the vague woolgathering of Bradley or the non-committal, tolerant hypothesising of the later Carnap as in the voluptuous revelling in multiplicity of linguistic philosophers. And on the other point, it is hard to think of a less adequate description than 'evasive' for Professor Ryle and Mr Strawson. *The Concept of Mind*, it might well be said, is not evasive enough. Moore, Wittgenstein, and Austin were all, in their different ways, unsystematic philosophers: Moore by a kind of temperamental accident, Wittgenstein as a matter of theory, Austin through a methodological decision. One can agree with Mr Gellner that this is a defect while thinking that it is immensely outweighed by their positive qualities when they are compared with, for example, Lotze or Hermann Cohen or Alexander.

In general, then, Mr Gellner fails to upset commonsense naturalism, partly through inability to identify it clearly enough, partly through misrepresenting it as an unsupported dogma; his arguments against the linguistic philosopher's concern with ordinary language are weak in detail, and implicitly concede a justification for this concern which is never openly admitted, but nevertheless they have a genuine bearing on an emotion of uncritical reverence which has tended to develop from the concern; and, finally, his attack on the flight from theory, both as a principle and a practice, though well-founded, subverts only a peripheral and dispensable aspect of linguistic philosophy. On the whole his examination has three main deficiencies: its violence, its distortion and, oddly enough, its evasiveness. The violence does not matter very much; except perhaps tactically in that his book will evoke only annoyance in the people to whom it is addressed, for it is exclusively negative in intention and execution (unless, of course,

sinuous creatures that they are, they glide out of the line of fire and give themselves up to the pleasures of *Schadenfreude*). The distortion consists in his remorseless attention to philosophical methodology and his general neglect, except for brief, exemplary notice, of the concrete philosophical work of his victims. To take just one example, it is extraordinary that Mr Strawson's criticism of formal logic, one of the most solid achievements of recent Oxford philosophy, is not even mentioned. Mr Gellner's evasiveness lies in his resolute refusal to reveal his own philosophical opinions or to define what he regards as the proper method of philosophy in any but the vaguest imaginable terms. It is perhaps an Oxford sense of decorum that inhibits me from attempting a sociological, as well as a philosophical, *tu quoque*. Incidentally his sociology is empirically shaky: logical positivism was a philosophy for gentlemen, linguistic philosophy is a somewhat more solidly middle-class affair. It is also a much less monolithic business than Mr Gellner makes it out to be (this belief, like the diagram, half joke, half obsession, has a slightly paranoiac flavour). Only a minority of Oxford philosophers come anywhere near fitting his description and the intellectual tone of that university is not the sole responsibility of philosophers. The unremitting hostility of historians, for example (voiced in Professor Trevor-Roper's review of *Words and Things*) is a discordant element which conflicts with Mr Gellner's picture of undisturbed complacency. But still, linguistic philosophy has its faults and it is a good thing for them to be pointed out even in the exaggerated outlines of a caricature.

ANTHONY QUINTON

The Study of Man. The Lindsay Memorial Lectures 1958. By Michael Polanyi.

Routledge & Kegan Paul, London, 1959. Pp. 102. 7s. 6d.*

Dear Dr Wisdom,

I struggled with this book recently and the more I think over it, the less I feel happy with some of the basic concepts Polanyi develops, whereas I still admire some others.

On the one hand he is saying that every understanding is tacit with human beings *and* animals, and on the other hand that every comprehension is an understanding, a grasping of disjointed parts into a comprehensive whole. Now I don't think that even Köhler's apes *understood* what they did although they certainly constructed a whole in order to obtain the banana.

* Editorial footnote: What follows is a letter we received from Professor van Lennep explaining why he could not review the book. We considered that the letter itself constituted an interesting review and accordingly, after obtaining Professor van Lennep's permission, we are printing it as a review.

REVIEWS

His equalisation of intellectual knowledge with knowing how to do something (skill) is, as far as I can see, playing with the ambiguity of the word knowledge. The crux of all this probably lies in the first sentence of his book: 'Man's capacity to think is his most outstanding attribute'. I would say: his capacity to reflect (= self-awareness, selfconsciousness) is what distinguishes him from animals, because this is the ontological fact which is responsible for the possibility of using language, i.e. symbols of all kinds. Therefore although there is no understanding without comprehension, there is comprehension without understanding.

But Polanyi says (p. 20) the word which covers them all (i.e. all comprehending experience) is simply understanding.

Skills in the sense of Polanyi are tacit knowledge and it is true that skills are generally not explicit, but is it not just playing with words to speak of *tacit knowledge*? Is knowledge not understanding in a reflective way what we comprehend tacitly?

Polanyi says no, because the act of understanding what science has made explicit remains a tacit act that in itself cannot be made explicit.

Understanding, however, I would say is a comprehension of things that are disconnected from their biological importance to the subject, and this is reserved to the human being. Although we say that a bird 'knows' how to build a nest he certainly does not *understand* his own procedure. And I just don't understand what Polanyi means when saying that understanding is a valid *form* of knowledge. What other forms are there?

Now I consider Polanyi a great man, and I very much admired his book *Personal Knowledge*, but as far as I am concerned I would consider this booklet less clear and convincing than the first one.

And since I apparently do not understand clearly his theory of understanding, nor appreciate his conceptual framework in the first chapter (the other two chapters are much clearer and I am more agreed with them) I feel that I am not sufficiently at one with him to criticise him in a review and not competent enough. A true epistemologist and not a simple applied psychologist should be much more up to the standard to do this job, the more so since he considers his own theory a *theory of knowledge*.

Therefore I am afraid that I must abandon my original plan to review his *Study of Man*, with my most sincere apologies for my late decision.

D. J. VAN LENNEP

Social Theory and Social Structure. By Robert K. Merton.

Revised and enlarged edition. The Free Press, Glencoe, Illinois, 1957.
Pp. 645. \$7.50.

ROBERT K. MERTON is not merely a professor of Sociology, he is a symbol and a programme. The volume under review is an expanded version of a

REVIEWS

work which has already made its mark and probably had a formative influence on many of those working within American sociology. What is he a symbol of? Not a simple empiricism, though empiricism is an ingredient. The empiricism which is associated with his programme is not so much an anti-transcendentalism, as past empiricisms have been, but a kind of high-minded puritanism: a deliberate determination in the interests of a far off salvation not to wallow now in the pleasures of interesting thought. Thought is not excluded, by any means: but it is a severely self-conscious methodical thought, preoccupied with its own propriety. Here perhaps we find the Calvinist personality of sociological thought, willing (or anxious?) to postpone present cognitive gratification in the interests of ultimate salvation.

What lies at the root of this? The acute awareness of sociology as a kind of under-developed country amongst the sciences, and a determination to set this right and catch up. As in these poor countries, the long-term policies required to escape under-development may conflict with the need to do as much good as possible at once.

There is a well known schema for what under-developed countries need: an adequate infra-structure, and a class with creative and entrepreneurial spirit. Professor Merton tries to supply, or aid the supply, of both these requirements. In other words, the book contains attempts both at the furthering of social theory and actual instances of handling of sociological problems. His work on the clarification of the much favoured 'functionalist' approach is already well known. On the substantive side, philosophers of science will be particularly interested in his studies of the social background of science.

Whether this self-conscious, long-term, and deliberate approach really will bear fruit, I suppose only time can tell. Professor Merton is superior to some he would call his masters. His followers may easily be led into stilted mumbo-jumbo. But with his preoccupation that sociological work should become cumulative one can strongly sympathise. And what are the alternatives? A pure and rigorous empiricism in sociology misses a good deal. At the other extreme, *weltgeschichtliche Betrachtungen*, though by no means to be despised, do not easily lend themselves to fruitful and diffused teaching. Even if they did, what would one do with a really large output of them?

So, in the meantime, it is possible that Professor Merton's choice is a right one, if perhaps not the uniquely right one. It is a pity that the path of virtue should sometimes seem such a stony one.

ERNEST GELLNER

Science and the Creative Spirit. Ed. by Harcourt Brown.
University of Toronto Press, 1958. Pp. xxvii + 165

FIVE representatives of various fields of humanities have produced this set of essays destined to throw several distinct beams of light upon their common

theme, viz. an evaluation of the influence exerted by humanities and sciences upon each other. What gives the volume its particular interest is the attempt to examine this problem by a team of humanists taking deliberately the attitude of 'critical sympathy' for science with the view of promoting a better understanding between the two camps.

The introductory article by Harcourt Brown, Professor of French at the Brown University, deals with basic concepts and definitions. Humanities are defined as being traditionally associated with 'the study of the record mankind has set down in books'; they are emphatically contemplative whereas sciences are preponderantly analytical; their primary subject is 'the creator and the creative imagination of man', 'the domain of the muses'. The sciences might be admitted as a part of this domain, if taken in their totality and without distinction between basic and applied sciences. The objective impersonality of science is in contrast with the 'self-communicative' ways of the humanist. True achievements of art can stand alone and independent, while those of the sciences never can be more than parts of something, like 'bricks in a cathedral'. In spite of their influential position, the sciences so far have failed to 'persuade' the modern humanist; for in the last instance science tends to destroy human values, being essentially a 'calculation of forces'. Similar ideas of Bertrand Russell of 1920 are quoted in conclusion.

Karl W. Deutsch, Professor of Political Science at the Massachusetts Institute of Technology, treats the subject of the scientific and humanistic knowledge in the growth of civilisation. Preservation of integrity and wholeness in the artistic interpretation of reality is contrasted by the isolating methods of sciences; and the singularity of situations dealt with by man in the arts, by the 'repetitiveness' of those studied by science. J. B. Connant's idea of science being accumulative in contrast to the art, is accepted with the reservation that the latter, too, is accumulative in a sense. Similarities in the scientific and artistic minds and in their respective ways are stressed, and their interplay in the growth of our civilisation is discussed and illustrated.

F. E. Priestley, Professor of English at the University of Toronto, deals with the effect of scientific progress upon the imagination in English literature. 'Significant ordering of experience' is a fundamental feature common to the arts and sciences, in spite of the differences in the working methods. Scientific influences in English literature are followed down to the beginnings of the Royal Society, with Jonathan Swift as the representative critic of the sciences of that epoch. The periods that follow are equally characterised by an interesting sequence of names, up to Charles Darwin's outstanding impact.

Harcourt Brown's article on the effect of sciences upon the development of French literature represents Blaise Pascal as the main figure characterising the initial period. Important English influences appear to have acted upon

REVIEWS

the relationship between sciences and arts in France; but the course the process took there was much different from the cognate process in England.

The concluding essay is written by a philosopher, Professor David Hawkins of the University of Colorado. The modern humanists have failed so far to understand the spirit of science, although every field of human creativity, especially that of philosophy, has been strongly influenced by science since the Renaissance. Connant's distinctive criteria of the artistic and scientific thinking cannot be accepted in their sharp formulation. Science is recognised as suitable to be placed 'alongside, if not among, the humanities'. A refreshing flash-history of the development of the scientific spirit and its influence illustrates these arguments, again with particular stress laid on the integrative and disintegrative ways of thought.

The papers have been trimmed so as to avoid major discrepancies in the individual outlooks, but they still have preserved strongly personal characters. The challenge of the book is so much the greater. The gaps are unavoidable in any work of this kind, and are obvious, being another added challenge. A scientist will be stimulated to much thought and comment on every page, and will close the book with a regret that the limitation of the study to the English and French area did not allow discussing such outstanding figures as Leonardo de Vinci and Goethe, both highly relevant for the general theme. Goethe's *Faust* contains many an invective against the Cartesian philosophy; Goethe's scientific contributions still are valued; and his interest in Purkynje and other 'heautognostic' physiologists might be considered as physiognomical for the subject. Another topic which would interest the scientist and which is only indirectly touched in the volume, is the influence of sciences upon art apart from literature, e.g. painting, sculpture, architecture, and music.

If this sample of opinions is representative of what the humanists nowadays feel about science, the impression is that of a defensive, or at least embarrassed attitude. Surprisingly little is said by the authors about the general moral problems connected with the growth of science. The voice of artists and humanists in this matter of undeniable urgency should be heard as a strong second fiddle to the first of the contemporary science.

JAN BELEHRADEK

Evidence and Inference. Edited by D. Lerner.

Free Press, Glencoe, Illinois, 1959. Pp. 164. \$4.00.

A SYMPOSIUM of such generality as this makes a very tantalising book. There are many extremely interesting suggestions in the various contributions, but the symposiasts have had to leave them undeveloped. Three contributions are worth singling out for particular mention; Raymond Aaron's remarks

REVIEWS

on the nature of historical and sociological 'individuals' (in the logical sense); Henry Hart and McNaughton's discussion of the law, particularly in view of some recent attempts to draw parallels between inductive inferences and legal verdicts; Erikson's account of the clinical method in psychiatry.

Aaron argues that while for historians individual human beings are the atoms, they are not always the units of the discourse. Units depend upon the interaction of several factors. For instance a battle is a unity for some historical purposes, because it is *both* spatio-temporally located and organised by the interaction of the plans of the commanders. However when we come to such historical units as 'periods' and 'cultures' we have no such clearly objectifiable individuating devices. What is chosen as the individual is more and more arbitrary the more diffuse it becomes. The formation of such individuals is a useful schematism but 'the error originates when, departing from the facts which prove the diversity and equivocal reality of *ensembles*, one infers a kind of metaphysics which transforms these *ensembles* into living beings, fated to be born and to die'.

In discussing the operation of the law Hart and McNaughton emphasise the peculiarity of the legal treatment of fact. The 'most distinctive practices of the law' they take to be found in the 'formal and official settlement of a controversy'. That is, as they put it, when a case reaches the courts, 'victory, and not accommodation is the object of the parties'. So that truth is what is left over from a process of legal attrition, rather than the total set of propositions established beyond reasonable doubt.

The discussion of clinical psychology also emphasises the difference between this kind of activity and science proper. Here, the mutual interaction of patient and therapist is the essence of the curative process. Nothing in the way of general law is invoked in what turns out to be a unitary process of diagnosis and treatment.

There are also contributions on physics (M. Deutsch), sociology (P. Lazarfield), and physiology (J. Fine). Each of these contains interesting, suggestive but sketchy material. There is an introduction by D. Lerner, designed to show the contrast between totalitarian and empirical theories of science.

R. HARRE

RECENT PUBLICATIONS ON THE PHILOSOPHY OF SCIENCE

(a) BOOKS RECEIVED FOR REVIEW

- Beauregard, O. C. de, *Théorie synthétique de la relativité restreinte et des quanta*, Gauthier-Villars, Paris, 1957, pp. xii + 200, \$9.30.
- Beckner, M., *The Biological Way of Thought*, Oxford University Press, London, 1959, pp. viii + 200, 48s.
- Bergmann, G., *Meaning and Existence*, University of Wisconsin Press, 1960, pp. x + 273, \$1.75.
- Borsuk, K. and Szmielew, W., *Foundations of Geometry*, North-Holland Publishing Co., Amsterdam, 1960, pp. 456, 90s.
- Bouvére, K. L. de, *A Method in Proofs of Undefinability*, North-Holland Publishing Co., Amsterdam, 1959, pp. vii + 64, 15s.
- Bragg, Sir William, *The Universe of Light*, Dover Publications, New York; London, Constable & Co., 1960, pp. viii + 283, \$1.85.
- Brogie, L. de, *Non-Linear Wave Mechanics*, Elsevier Publishing Co., 1960, pp. xii + 304, 57s.
- Chandrasekhar, S., *Radiative Transfer*, Dover Publications, New York; London, Constable & Co., 1960, pp. xiv + 393, \$2.25.
- Curry, H. B. and Feys, R., *Combinatory Logic*, North-Holland Publishing Co., Amsterdam, 1958, pp. xvi + 417.
- Danto, A. and Morgenbesser, S., *Philosophy of Science*, Meridian Books, New York, 1960, pp. 477, \$1.65.
- Evans, H. M. (Ed.), *Men and Moments in the History of Science*, University of Washington Press, 1959, pp. viii + 226, \$4.50.
- Farber, M., *Naturalism and Subjectivism*, Thomas, Illinois, 1959, pp. xvi + 389.
- Fowler, J. M., *Fallout. A Study of superbombs, strontium 90 and survival*, Basic Books, New York, 1960, pp. 235, \$5.50.
- Freudenthal, H., *Lincos, Design of a Language for Cosmic Intercourse*, Part I, North-Holland Publishing Co., Amsterdam, 1960, pp. 224.
- Fritz, E., *Energies of the Universe*, Philosophical Library, New York, 1960, pp. 124, \$4.75.
- Gibson, Q., *The Logic of Social Enquiry*, Routledge & Kegan Paul, London, 1960, 24s.
- Gillispie, C. C., *The Edge of Objectivity*, Princeton University Press; London, Oxford University Press, 1960, pp. vii + 562, 42s.
- Gonseth, F., *La Métaphysique et l'ouverture à l'expérience*, Presses Universitaires de France, Paris, 1960, pp. 291, NF 12.
- Hall, A. R. (Ed.), *The Making of Modern Science*, Leicester University Press, 1960, pp. 54, 6s.
- Harré, R., *An Introduction to the Logic of the Sciences*, Macmillan, London, 1960, pp. viii + 180, 21s.
- Hoskin, M., *William Herschel*, Billing, London, 1959, pp. 48, 2s. 6d.
- Isaacs, N., *New Light on Children's Ideas of Number*, E.S.A., London, 1960, pp. 40, 3s. 6d.

RECENT PUBLICATIONS

- Koninck, C. de, *The Hollow Universe*, Oxford University Press, London, 1960, pp. xii + 127, 12s. 6d.
- Kraft, V., *Problems der Wissenschaftstheorie*, Springer-Verlag, Vienna, 1960, pp. vi + 266.
- Landé, A., *From Dualism to Unity in Quantum Physics*, Cambridge University Press, London, 1960, pp. xv + 114, 18s. 6d.
- Levitt, M., *Freud and Dewey on the Nature of Man*, Philosophical Library, New York, 1960, pp. 180, \$3.75.
- Madden, E. H. (Ed.), *Theories of Scientific Method: The Renaissance through the Nineteenth Century*, University of Washington Press, 1960, pp. iv + 346, \$6.50.
- Madden, E. H., *The Structure of Scientific Thought*, Routledge & Kegan Paul, London, 1960, pp. ix + 381, 35s.
- Melsen, A. G. van, *From Atomos to Atom* (trans. by H. J. Koren), Duquesne University Press, 1952, pp. xii + 240, \$4.25.
- Mises, L. von, *Epistemological Problems of Economics*, Van Nostrand, London, 1960, pp. xxiii + 239, 41s. 6d.
- Mudry, J., *Philosophy of Atomic Physics*, Philosophical Library, New York, 1958, pp. 136, \$3.75.
- Mukerjee, R., *The Symbolic Life of Man*, Hind Kitabs, Bombay, 1959, pp. xii + 294, Rs. 15.00.
- Mulckhuysen, J. J., *Molecules and Models: Investigations on the Axiomatization of Structure Theory in Chemistry*, Rototype, Amsterdam, 1960, pp. v + 67.
- Nidditch, P. H., *Elementary Logic of Science and Mathematics*, University Tutorial Press, London, 1960, pp. vii + 369, 18s.
- Palter, R. M., *Whitehead's Philosophy of Science*, University of Chicago Press; London, Cambridge University Press, 1960, pp. xv + 248, 60s.
- Pike, K. L., *Language in Relation to a Unified Theory of the Structure of Human Behaviour*, Summer Institute of Linguistics, California, 1960, pp. vii + 146, \$3.50.
- Rosen, E. (Trans. with Introduction), *Three Copernican Treatises*, 2nd edn., Dover Publications, New York; London, Constable & Co., 1959, \$1.75.
- Rougier, L., *La Métaphysique et le Langage*, Flammarion, Paris, 1960, pp. 247, NF 9.50.
- Schilpp, R. A. (Ed.), *The Philosophy of C. D. Broad*, Tudor Publishing Co., New York; London, Cambridge University Press, 1960, pp. xii + 866, 110s.
- Schütte, K., *Beweistheorie*, Springer-Verlag, Vienna, 1960, pp. x + 353.
- Sidman, M., *Tactics of Scientific Research*, Basic Books, New York, 1960, pp. x + 428, \$7.50.
- Simons, J. H., *A Structure of Science*, Philosophical Library, New York, 1960, pp. v + 269, \$4.75.
- Stallo, J. B., *The Concepts and Theories of Modern Physics*, Harvard University Press; London, Oxford University Press, 1960, pp. xxix + 325, 38s.
- Suppes, P., *Axiomatic Set Theory*, Van Nostrand Co., London, 1960, 45s.
- Suppes, P., Henkin, L. and Tarski, A. (Eds.), *The Axiomatic Method with special reference to Geometry and Physics*, North-Holland Publishing Co., Amsterdam, 1959, pp. xi + 488, 90 guilders.
- Swanson, M. A., *Scientific Epistemologic Backgrounds of General Semantics*, Institute of General Semantics, 1959, pp. vii + 81.
- The Royal Society, *The Correspondence of Isaac Newton II, 1676-86*, Cambridge University Press, London, 1960, pp. xii + 552, 147s.

RECENT PUBLICATIONS

- Urmson, J. O. (Ed.), *The Concise Encyclopaedia of Western Philosophy and Philosophers*, Hutchinson, 1960, pp. 431, 50s.
- Vigier, J.-P., *Structure des micro-objets dans l'interprétation causale de la théorie des Quanta*, Gauthier-Villars, Paris, 1956, pp. xi + 192, \$8.70.
- Vuillemin, J., *Mathématiques et Métaphysique chez Descartes*, Presses Universitaires de France, Paris, 1960, pp. 188, NF 16.
- Weisheipl, J. A., *The Development of Physical Theory in the Middle Ages*, Sheed & Ward, London, 1960, pp. 92, 4s.
- Whitehouse, W. A., *Order, Goodness, Glory*, Oxford University Press, London, 1960, pp. 83, 9s. 6d.
- Yourgrau, W. & Mandelstam, S., *Variation Principles in Dynamics and Quantum Theory*, 2nd edn., Pitman, London, 1960, pp. xi + 180, 32s. 6d.

(b) ARTICLES

- Abetti, G., 'Le scoperte astronomiche di Galileo', *Scientia*, 1960, **95**, 77-85
- Apostel, L., 'Sur la méthode en théorie de la connaissance', *Revue Internationale de Philosophie*, 1959, **13**, 460
- Barker, S. F. and Achinstein, P., 'On the New Riddle of Induction', *The Philosophical Review*, 1960, **49**, 511-522
- Bensmer, J. and Vidich, A., 'Social Theory in Field Research', *American Journal of Sociology*, 1960, **65**, 577-584
- Bergmann, G., 'Ineffability, Ontology, and Method', *The Philosophical Review*, 1960, **49**, 18-41
- Bierstadt, R., 'Sociology and Humane Learning', *American Sociological Review*, 1960, **25**, 3-9
- Black, M., 'Linguistic Relativity: The Views of Benjamin Lee Whorf', *The Philosophical Review*, 1959, **48**, 228-238
- Black, M., 'Possibility', *The Journal of Philosophy*, 1960, **57**, 117-126
- Blalock, H. M., Jr., 'Correlational Analyses and Causal Inferences', *American Anthropologist*, 1960, **62**, 624-631
- Bourdier, F., 'Trois siècles d'hypothèses sur l'origine et la transformation des êtres vivants (1550-1859)', *Revue D'Histoire des Sciences*, 1960, **13**, 1-44
- Builder, G., 'The Lorentz Transformations', *The Australian Journal of Physics*, 1959, **12**, 300-303
- Builder, G., 'The Fundamental Concepts of Relativity Theory', *The Australian Journal of Science*, 1959, **22**, 87-97

ERRATA

In the article-review 'Wittgenstein's Theory and Practice of Philosophy' by G. Kreisel, this *Journal*, 1960, **11**, the following two corrections should be made:

- page 242, note 3 for 'Kleene and Mostowski' read 'G. Kreisel'.
- page 251, note 1 in the last line, for 'note 2' read 'note 3'.